

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 05:51:34 ; Search time 8358 Seconds
(without alignments)
11491.431 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ttggcattgcgcgcagcaaca.....tngagagaccctcccaacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2024	99.7	2031	6	AX082595 Sequence
2	1655.2	81.5	5680	9	AB018271 Homo sapi
3	1652	81.3	16282	9	AF400226 Homo sapi
4	1637.6	80.6	3307	6	AX834329 Sequence
5	1637.6	80.6	3307	6	AK096713 Homo sapi
6	1637.6	80.6	4112	9	AK128632 Homo sapi
7	1334.6	65.2	7194	6	CQ732823 Sequence
8	1190.8	58.6	17206	10	AF396878 Mus muscu
9	1189.2	58.6	22218	10	AF396879 Mus muscu
10	1022	50.3	1022	9	AF083131 Homo sapi
11	674.4	33.2	15293	6	CQ729862 Sequence
12	655.4	32.3	6007	6	AX677291 Sequence
13	655.4	32.3	14491	9	AB007934 Homo sapi
14	655.4	32.3	17518	9	AB029290 Homo sapi
15	655.4	32.3	17734	9	AF141968 Homo sapi
16	655.4	32.3	19153	9	AF317696 Homo sapi
17	651	32.1	17333	10	AF150755 Mus muscu
18	541.8	26.7	5630	10	AK122291 Mus muscu
19	438.8	21.6	547	6	CQ429811 Sequence

20	422.8	20.8	531	6	CO420937 Sequence
21	357	17.6	401	6	AX071203 Sequence
22	350	17.2	5523	10	AK129201 Mus muscu
23	266.8	13.1	4495	6	AK38861 Sequence
24	262.8	12.9	344	6	AX307126 Sequence
25	250.2	12.3	278	6	AX920270 Sequence
26	250.2	12.3	278	6	BD055803 Sequence
27	234.4	11.5	55703	9	AL512448 Human DNA
28	181	8.9	22604	2	AC094308 Rattus no
29	181	8.9	249618	2	AC136131 Rattus no
30	177.8	8.8	3684	6	AX713588 Sequence
31	177.8	8.8	3684	9	AK055189 Homo sapi
32	177.2	8.7	146080	10	AC124386 Mus muscu
33	165.2	8.1	209299	5	AC145916 Gallus ga
34	162.2	8.0	12989	3	DMGROVIN
35	162.2	8.0	12995	6	CO600345 Sequence
36	162.2	8.0	16798	3	AF200425 Drosophila
37	162.2	8.0	17678	6	CO598494 Sequence
38	162.2	8.0	18011	6	CO600516 Sequence
39	162.2	8.0	27845	6	CO600369 Sequence
40	153.6	7.6	388	6	BD072321 Secreted
41	136	6.7	67033	9	AL137008 Human DNA
42	115.8	5.7	21491	9	AF325340 Homo sapi
43	115.8	5.7	113853	9	AL137853 Human DNA
44	115.8	5.7	173077	2	AC024511 Homo sapi
45	115.8	5.7	227801	2	AL713993 Homo sapi

ALIGNMENTS

RESULT 1
AX082595
LOCUS AX082595 Sequence 46 from Patent WO011047.
DEFINITION AX082595
ACCESSION AX082595
VERSION AX082595.1 GI:13184697
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Bowman, B.M. and Wang, K.
TITLE Dna sequences isolated from human colonic epithelial cells
JOURNAL Patent: WO 011047-A 46 15-FEB-2001;
Bayer Corporation (US)
FEATURES
SOURCE location/Qualifiers
1..2031
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

misc_feature 1162
/note="Unknown"
misc_feature 1202
/note="Unknown"
misc_feature 1229
/note="Unknown"
misc_feature 1277
/note="Unknown"
misc_feature 1885
/note="Unknown"
misc_feature 1889
/note="Unknown"
misc_feature 2013
/note="Unknown"

ORIGIN

Query Match 99.7%; Score 2024; DB 6; Length 2031;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2031; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GY 1 TTGGCATTGCTGACGAACACATCATCTCAGCTCCGCCAGCCCTTGATA 60

Db 1 TTGCCATGCTGACGAAACCAATCAATCATCTCTCAGCTCCGCCCCAGCCCTTGAAATA 60
Qy 61 TGAATCTTAAGGACAGACAGAGAAACATCGGCAATCGCGTGAAGTGTACTGTAACA 120
Db 61 TGAATCTTAAGGACAGACAGAGAAACATCGGCAATCGCGTGAAGTGTACTGTAACA 120
Qy 121 CAAGCTCATATATGATTAAGTGAACAAACCTGGCCACAGTACTGAATGAGCCCTGGGAA 180
Db 121 CAAGCTCATATATGATTAAGTGAACAAACCTGGCCACAGTACTGAATGAGCCCTGGGAA 180
Qy 181 GGCCTTTCTATCCAAGGAAGTATGTGGCAGCCGACACCTTTTACAGTCACTTAAAGAA 240
Db 181 GGCCTTTCTATCCAAGGAAGTATGTGGCAGCCGACACCTTTTACAGTCACTTAAAGAA 240
Qy 241 GATGTCAAAAAGGCTGTGGCAGCTGGATGAAGCCATTTCTCAATCAATCACTGATTCAT 300
Db 241 GATGTCAAAAAGGCTGTGGCAGCTGGATGAAGCCATTTCTCAATCAATCACTGATTCAT 300
Qy 301 GACAAATATGATCAGATCCTTGAGAGCCTGGAACGCATCGTGAACGCTGAGGACCA 360
Db 301 GACAAATATGATCAGATCCTTGAGAGCCTGGAACGCATCGTGAACGCTGAGGACCA 360
Qy 361 CCTCTATCTCTGCAAGGTTGAGAAAGATCAAGAAACAGATCACTGAAAATTAAGATGTG 420
Db 361 CCTCTATCTCTGCAAGGTTGAGAAAGATCAAGAAACAGATCACTGAAAATTAAGATGTG 420
Qy 421 TCAGTGAACAATGGAAGGCTACAGCCGTTGATGAACCTGTTAAACAGAGGGAGAGAA 480
Db 421 TCAGTGAACAATGGAAGGCTACAGCCGTTGATGAACCTGTTAAACAGAGGGAGAGAA 480
Qy 481 ATGATTTAGATCTGGGGGGAATGAAGATCAATCTGCAAGGCTGTTAGATGAAG 540
Db 481 ATGATTTAGATCTGGGGGGAATGAAGATCAATCTGCAAGGCTGTTAGATGAAG 540
Qy 541 CTTGACCAAAATGGTTTCAATTTGGAGAACATACACACTGCTGGAAGAGAGAGGCC 600
Db 541 CTTGACCAAAATGGTTTCAATTTGGAGAACATACACACTGCTGGAAGAGAGAGGCC 600
Qy 601 AAACCTCTGGAATGATGAGCTAGACAGAAAGTTTGGTGTATCAATGATGATGATA 660
Db 601 AAACCTCTGGAATGATGAGCTAGACAGAAAGTTTGGTGTATCAATGATGATGATA 660
Qy 661 GTTACCATTAAGATATCTCAAGATTTCAATCGGGAACCTGGAAGATCCTGGAATTGATCCT 720
Db 661 GTTACCATTAAGATATCTCAAGATTTCAATCGGGAACCTGGAAGATCCTGGAATTGATCCT 720
Qy 721 TCAGTATGAACCAACAGCAAGCAAGCAGAGACCAATAAGGAAGAAATGATGACATA 780
Db 721 TCAGTATGAACCAACAGCAAGCAAGCAGAGACCAATAAGGAAGAAATGATGACATA 780
Qy 781 CAGGAGAGCTGATATATGATTAACCTAGGCTTCAACTCATTCGCGGATGTGGAGAG 840
Db 781 CAGGAGAGCTGATATATGATTAACCTAGGCTTCAACTCATTCGCGGATGTGGAGAG 840
Qy 841 CCTGATAAACCAATGCTCAAGAGATATCATGAGTAAATTCAGGATGGATCTCT 900
Db 841 CCTGATAAACCAATGCTCAAGAGATATCATGAGTAAATTCAGGATGGATCTCT 900
Qy 901 AAATAAAGCTTGAAGAAGCCGATGTAACCACTTGAAGCAATGACGCTGCTCAGTAC 960
Db 901 AAATAAAGCTTGAAGAAGCCGATGTAACCACTTGAAGCAATGACGCTGCTCAGTAC 960
Qy 961 AGATGACCTGACAGGCGTATTTTCTGGGATGATATTCAGAGGTGAAGTTCGCTTCAAT 1020
Db 961 AGATGACCTGACAGGCGTATTTTCTGGGATGATATTCAGAGGTGAAGTTCGCTTCAAT 1020
Qy 1021 GTCTCAATTTGAACAGCATCTGAAATCTGCAAGCAGCATGTTGCTAGACCAAT 1080
Db 1021 GTCTCAATTTGAACAGCATCTGAAATCTGCAAGCAGCATGTTGCTAGACCAAT 1080
Qy 1081 TTAAGTCTGAGCCTTATCAACAGCAGATGAGTGTACGATCACTCAGCAGACGC 1140
Db 1081 TTAAGTCTGAGCCTTATCAACAGCAGATGAGTGTACGATCACTCAGCAGACGC 1140

Db 1081 TTAAGTCTGAGCCTTATCAACAGCAGATGAGTGTACGATCACTCAGCAGACGC 1140
Qy 1141 TTTTGTAAAGAAAGTAAAGAGAGAGTGAACAAACACTGTTCTAGACCCAAATAATG 1200
Db 1141 TTTTGTAAAGAAAGTAAAGAGAGAGTGAACAAACACTGTTCTAGACCCAAATAATG 1200
Qy 1201 ANCTGATATGATATGCTAGGCTGAGAGAGAAATCATCAACAGACATTAACCTGGA 1260
Db 1201 ANCTGATATGATATGCTAGGCTGAGAGAGAAATCATCAACAGACATTAACCTGGA 1260
Qy 1261 GGGTGTCTATTAAGCCATGGGTGATGTTCCAAACATGCTGTATGAGCTCTGATCTGA 1320
Db 1261 GGGTGTCTATTAAGCCATGGGTGATGTTCCAAACATGCTGTATGAGCTCTGATCTGA 1320
Qy 1321 CACACGCGAGGGCTTGAATGAGCAGAAACCTGTTGAGAGAGACCTTAAGCCATTTG 1380
Db 1321 CACACGCGAGGGCTTGAATGAGCAGAAACCTGTTGAGAGAGACCTTAAGCCATTTG 1380
Qy 1381 AAATTGAACTTGCCAAAGCATCATGTGCTCAAAATGATGATTAAGCCATCAGTCCACAG 1440
Db 1381 AAATTGAACTTGCCAAAGCATCATGTGCTCAAAATGATGATTAAGCCATCAGTCCACAG 1440
Qy 1441 TGGAGCCGTTAATTAAGCAGAAATGATCTAATGAAATCAAGTGCAGAGAGAAAGCAA 1500
Db 1441 TGGAGCCGTTAATTAAGCAGAAATGATCTAATGAAATCAAGTGCAGAGAGAAAGCAA 1500
Qy 1501 GCAACCTTGAGAAACAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTTGGAAAAA 1560
Db 1501 GCAACCTTGAGAAACAGCTAGAGGTTTAAATCAACGCTGGGCAAAATGTTTGGAAAAA 1560
Qy 1561 CAGAACAAAGAGAGCAGCAGCTGATGATGCTGCTGGCCAGGCAAAAGGTTTCAATGGG 1620
Db 1561 CAGAACAAAGAGAGCAGCAGCTGATGATGCTGCTGGCCAGGCAAAAGGTTTCAATGGG 1620
Qy 1621 AAATTGAGATTTGACAGCAGTGGCTGATGACACGAGGCTATCTGTTGGCATCTAAAC 1680
Db 1621 AAATTGAGATTTGACAGCAGTGGCTGATGACACGAGGCTATCTGTTGGCATCTAAAC 1680
Qy 1681 CGCTGGAGGTTTAAACCGGAAAGAGCAAGCAGACAGTTAATGCCATATGSAAGTCTGG 1740
Db 1681 CGCTGGAGGTTTAAACCGGAAAGAGCAAGCAGACAGTTAATGCCATATGSAAGTCTGG 1740
Qy 1741 CTGCTTTGAACTAAAGAGAAACATTAAGTCTGATGACAGAAAGGCCAGCAGATGC 1800
Db 1741 CTGCTTTGAACTAAAGAGAAACATTAAGTCTGATGACAGAAAGGCCAGCAGATGC 1800
Qy 1801 TTGCAAGATGCCCAAAATCTGCAAGACAAATATTTGACCAAGACATTAATTAACCTGAAAA 1860
Db 1801 TTGCAAGATGCCCAAAATCTGCAAGACAAATATTTGACCAAGACATTAATTAACCTGAAAA 1860
Qy 1861 AAAAAAAGGCTTGAAGAAGCCGATGTAACCAAGGCTGAAATTCAGAGGCGCGCTG 1920
Db 1861 AAAAAAAGGCTTGAAGAAGCCGATGTAACCAAGGCTGAAATTCAGAGGCGCGCTG 1920
Qy 1921 AATTCTAGACCTGCGGCGCGCGCTGAGACCCCTTAATGATGATGATGATGATGATGAA 1980
Db 1921 AATTCTAGACCTGCGGCGCGCGCTGAGACCCCTTAATGATGATGATGATGATGATGAA 1980
Qy 1981 TCACTAGTGGCGGCTGCAAGTTCGACATATNGAGAGAGCTCCCAACGCGT 2031
Db 1981 TCACTAGTGGCGGCTGCAAGTTCGACATATNGAGAGAGCTCCCAACGCGT 2031

RESULT 2
AB018271
LOCUS AB018271 5680 bp mRNA linear PRI 10-MAY-2002
DEFINITION Homo sapiens mRNA for KIAA0728 protein, partial cds.
ACCESSION AB018271
VERSION AB018271.2 GI:20521140
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
JOURNAL
COMMENT
FEATURES
SOURCE

1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.
Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
DNA Res. 5 (5), 277-286 (1998)
9872452
2 (bases 1 to 5680)
Ohara, O., Suyama, M., Nagase, T., Ishikawa, K. and Kikuno, R.
Direct Submission
Submitted (08-OCT-1998) Oasamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yana 1532-3, Kiseazsu, Chiba 292-0612, Japan (E-mail: cdna@info.kazusa.or.jp, Tel: +81-438-52-3913, Fax: +81-438-52-3914)
On May 9, 2002 this sequence version replaced gi:3882176.
Location/Qualifiers
1. 5680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="f122477"
/sex="male"
/tissue_type="brain"
/clone_lib="pbluescriptII SK plus"
/dev_stage="adult"
/note="This sequence was replaced that of hK03530 cDNA as a representative cDNA sequence for KIAA0728."
1. 5680
/gene="KIAA0728"
1. 4914
/gene="KIAA0728"
/codon_start=1
/product="KIAA0728 protein"
/protein_id="BAA34448.2"
/db_xref="GI:20521141"
/translation="EELMPWLTETOSIIISQIPALAEYETIRQOEHRRLRELIASH
KPHIDKNKTPQQLLESPGEGFSIQEYVAADLYSQIKEDVKRVALDEALISQST
QFHDKIDQIIIESLEIRIVERLQPSISAEVEKIQIISENVSVDMEKQLPLETK
QGRSEMIARSGGTGKDIISAKAVQDKDQWYIEMNIHLVVERSAVKLDDWETK
CDHMSLVITIKTDQDFRLEDDPGIDPSVVOOQAAETREIIBEDYIOBELDIYNG
SELIACGEPDPIVYKSIDELINAMPSINKWMDRIKLEMAQAAVQDQGAQNP
DMVDIAGCKLASNPITGDTETVQKQIBELKQSEAYQQQIENRNLHQBELIKY
TESDKRTVDDPLMKLEIMDSLEIRIINRQKLEGALLAGQPHALDELAVLTHT
EGLSEQRVPVGDPAKLEIELAKHVLQNDVLAHQSTVEAVNKAQNDLISAGEBAS
NLQNKLEVLNQRQNVLEKTEQRKQDGLRQAKGPHGEIEDLQWLTDERHLLAS
KPLGGLPETAKLEOLNMEVCAAFEAKEBETKYSIMQGOQWLACPSAETNIDODIN
NIKEKMSVETKIKNERKLEALNINLAMEPHNSIQDPIINMTQAEQRLVNASRSLI
DTVLFOIDENKVPANESVNSHREQIIEIDKGTGTHLYFSQKQDVPLIKNLLISVSRNE
KVQVRLIVERGSLDDARKRAKQFHEAWSKLEWLESEKSDLELAINPDKIKTDL
AOKHEFOKSLGAKHSVYDITNRGSLKEKTSIADNDLKDDMLSELKDWDICGS
VERONKLEALPFSQFDALQALIDWLYRVEPOLAEDQPYHGIDLVNMLIDHKAF
QKELGRTSSVQALKRSARELIEGSRDSSVYQWOMELSTRMTWELISKQTRLS
AALRQAEERHSVYVALLLEALAEQTLRFHVLPDDEDAITLIDQKPEKTSERK
AEUNKATITGDTVLAICHPDSITTKMTITIRAFBEVLMAWQHQRLASALAGLI
AKOELLEALLAMLQMAETTLTDKQEVIPQIEBVKALIEHQTFMEBMRKQADLVK
VTKYKRRADPSSLOSHIPLYLDKRAKRFPSASLYPSGSOQIETKQPRVLLVS
KMOQVLLALERRKLDALDRLELREFANFDIWRKKYMRMKNKSSVMDFFR
IDKQDQKITRROEFDIGLISKPSPTSLRLEMSAVADIDRDQDGYDIYVEPALHPK
DAYKPIIDAKTIDBEVTRQVAKKCAKRFQYEQJGDKNYRFFLQNGDSQQLVYAI
LRSTVTRVGGAMALDEFLVKNPCBARKGTNNLEKFLIADGASQGMARFRPR
RSRPSRSGASNRSLTSVSSQAAQASPOVATTPPKLHLPTRYGKFWMLNVSMSIP
CKAECSDPVPASAEPTPIQGSKLRLPGYLSGKGFHSGBSGITTTAAARVTPFAS
KTPSRGSRGSRGASGRASRGRSDSDPDISIIGVCSQVETVPTQHRTPRAGSR
PSTAKPKIPRPQRKSPASKLIDKSKR"

ORIGIN

Query Match 81.5%; Score 1655.2; DB 9; Length 5680;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1804; Conservative 0; Mismatches 73; Indels 18; Gaps 9;

Oy 8 GCTGACAGAAACAAATCAATCATCTCTCAGC--TCCGCCCCAGCCCTTGAATATGATAC 66
Db 18 GCTGACAGAAACAAATCAATCATCTCTCAGCTTCCGCCCCAGCCCTTGAATATGAAAC 77
Oy 67 TCTAAGCAGCAGCAGGAAGAAACATCTGCACATCTGCGTAGTTGATAGCTGAACCAAGCC 126
Db 78 TCTAAGCAGCAGCAGGAAGAAACATCTGCACATCTGCGTAGTTGATAGCTGAACCAAGCC 137
Oy 127 TCATATGATATGATGAACAAACTGG----CCACAGTACTGAATGAGCCCTGGGGAAAG 182
Db 138 ACATATGATATGATGAACAAACTGGGGCAAGTTACTGGAATGAGCCCTGGGGAAAG 197
Oy 183 CTTTTCATCAAGAGAGAGATATGAGGACCGACACCTTTACAGTACATTAAGAAGA 242
Db 198 CTTTTCATCAAGAGAGATATGAGGACCGACACCTTTACAGTACATTAAGAAGA 257
Oy 243 TGTCAAAAGCGTCTGTGCACTGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 302
Db 258 TGTCAAAAGCGTCTGTGCACTGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 317
Oy 303 CAAGATGATCAGATCTTGAAGCCTGGAACCGATCTGTGAACCTGTGAGGCCACCC 362
Db 318 CAAGATGATCAGATCTTGAAGCCTGGAACCGATCTGTGAAGCCTGTGAGGCCACCC 377
Oy 363 CTTATCTCTGACAGAGTTGGAAGATCAAGGAACAGATCAGTGAATAAATGATGTCTC 422
Db 378 CTTATCTCTGACAGAGTTGGAAGATCAAGGAACAGATCAGTGAATAAATGATGTCTC 437
Oy 423 AGTAGACATGAAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 482
Db 438 AGTAGACATGAAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 497
Oy 483 GATTGCTAGATCTGGGGGGGACTGATTAAGACATATCTGCCAAGCTGTTCAGATTAAGCT 542
Db 498 GATTGCTAGATCTGGGGGGGACTGATTAAGACATATCTGCCAAGCTGTTCAGATTAAGCT 557
Oy 543 TGACCAATAGTTTCACTTTGGGGAACATACACACTGGTGGGAAGAGGAGGACCA 602
Db 558 TGACCAATAGTTTCACTTTGGGGAACATACACACTGGTGGGAAGAGGAGGACCA 617
Oy 603 ACTACGAGATGATGAGAGCTAGACAGAAAGTTCTGCTGTGATCAATGATTAAGT 662
Db 618 ACTACGAGATGATGAGAGCTAGACAGAAAGTTCTGCTGTGATCAATGATTAAGT 677
Oy 663 TACCATTAAGATATCTCAAGATTTCACTCGGAGCTTGAAGATCTTGAATTAAGTCTTC 722
Db 678 TACCATTAAGATATCTCAAGATTTCACTCGGAGCTTGAAGATCTTGAATTAAGTCTTC 737
Oy 723 AGTAGTAACCAACAGCAAGACAGACAGACCATTAAGGAAATTAAGTAGACTACA 782
Db 738 AGTAGTAACCAACAGCAAGACAGACAGACCATTAAGGAAATTAAGTAGACTACA 797
Oy 783 GAGAGAGCTGATATGATTAATTAAGTTCGAAGCTTGAAGCTATTCGCGCATGTGGGAGCC 842
Db 798 GAGAGAGCTGATATGATTAATTAAGTTCGAAGCTTGAAGCTATTCGCGCATGTGGGAGCC 857
Oy 843 TGATTAACCAATGTCGAAGAGATATACATGATTAATTAATTAAGTTCGATTA 902
Db 858 TGATTAACCAATGTCGAAGAGATATACATGATTAATTAATTAAGTTCGATTA 916
Oy 903 ATTAAGTTGGAAGACCGGATTAAGCAACTTGAAG-----AGCAATGACAGCTGCCGTC 956
Db 917 ATTAAGTTGGAAGACCGGATTAAGCAACTTGAAGCAAGTGAAGCTGCCGTCAGT 976
Oy 957 GTACAGATGACAGCAGCGCGTATTTTCTGGGTAGATATTAAGTGTGAAGTTCGCTT 1016
Db 977 ACCAGATGACAGCAGCGCGTATTTGAATGAGTATTAAGTGTGAAGTTCGCTT 1036
Oy 1017 CAATGCTCAATTTGAACAGCATCTGGAACCTGTCAGAGCAGATTAATGTCAGAG 1076
Db 1037 CAATGCTCAATTTGAACAG-ATCTGGAACCTGTCAGAGCAGATTAATGTCAGAG 1095
Oy 1077 CAATTTAAGTCTGAGGCTATCAACAGCAATAG-ACTGATGACAGTCACTCAACGCA 1135

Db	1096	CAATTAAAGCTGAGGGCTTCAACAGCAGATGAAATGAAAGACTGATCATCAAGCA	1155
Qy	1136	GACGCTTTTCTTAAGAAATGAACAGNAGAGTGAACAACACACTGTTAGCCCAAA	1195
Db	1156	GA-GCTTTTCTTAAGAAATGAACAGNAGAGTGAACAACACACTGTTAGCCCAAT	1214
Qy	1196	AATGAGCTGATATGTATAT-GGCTAGCTTGAAGNAGATCATCAACAGACATTA	1254
Db	1215	AATGAACTGAATATGTATATGTATGAGTCTGGAAGAGATCATCAACAGACATTA	1274
Qy	1255	ACTGAGGGTGTCTTATAGCCNTGGTCACTTCAACATGCCCTGTATAGCTCTG--	1312
Db	1275	ACTGAGGGTGTCTTATAGCCNTGGTCACTTCAACATGCCCTGTATAGCTCTGCG	1334
Qy	1313	CATGTCGACACACACGAGGGCTTGTAGTGAAGAAACCTGTTGAGAGACCCCTTA	1372
Db	1335	ATGGCTGACACACACGAGGGCTTGTAGTGAAGAAACCTGTTGAGAGACCCCTTA	1394
Qy	1373	AGCCATTGAAATTTGAATCTTCCAGCATCATGTGCTCCAAATGATGTATTAGCCATCA	1432
Db	1395	AGCCATTGAAATTTGAATCTTCCAGCATCATGTGCTCCAAATGATGTATTAGCCATCA	1454
Qy	1433	GTCACAGTGAAGCCGTTATATTAAGCAGAAATGATCTTAATTGATCAAGTGAAGAGA	1492
Db	1455	GTCACAGTGAAGCCGTTATATTAAGCAGAAATGATCTTAATTGATCAAGTGAAGAGA	1514
Qy	1493	AGAAGCAGCAGCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTT	1552
Db	1515	AGAAGCAGCAGCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTT	1574
Qy	1553	GGAAGAAACAGAACAAAGAGAGCAGCTGATGTGCTTCCGCGCAGCCAAAGGCTT	1612
Db	1575	GGAAGAAACAGAACAAAGAGAGCAGCTGATGTGCTTCCGCGCAGCCAAAGGCTT	1634
Qy	1613	CGATGCGGAAATGAGGATTTGACAGCAGTGGCTGACAGACGAGGCTATCTGTTGGC	1672
Db	1635	CGATGCGGAAATGAGGATTTGACAGCAGTGGCTGACAGACGAGGCTATCTGTTGGC	1694
Qy	1673	ATCTTAAACCGCTGGAGGTTTACCGGAAACAGCAGGAGCAGCTTAATGCTCATTTGA	1732
Db	1695	ATCTTAAACCGCTGGAGGTTTACCGGAAACAGCAGGAGCAGCTTAATGCTCATTTGA	1754
Qy	1733	AGTCTGTGCTGCTTTGAAGCTTAAAGAAACATTAAGAGTCTGATGACAGAAAGCCCA	1792
Db	1755	AGTCTGTGCTGCTTTGAAGCTTAAAGAAACATTAAGAGTCTGATGACAGAAAGCCCA	1814
Qy	1793	GCGATGCTTGAAGATGCCCCAAATCTGACAGACAAATATTGACCAAGACATTAATTA	1852
Db	1815	GCGATGCTTGAAGATGCCCCAAATCTGACAGACAAATATTGACCAAGACATTAATTA	1874
Qy	1853	CTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1887
Db	1875	CTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1909

RESULT 3
AP400226
LOCUS
DEFINITION
Homo sapiens bullous pemphigoid antigen 1 ea (BPAG1) mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AP400226 16282 bp mRNA linear Pri 19-FEB-2002
Homo sapiens bullous pemphigoid antigen 1 ea (BPAG1) mRNA, complete cds.
AP400226 AP400226.1 GI:18157650
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Okumura, M., Yamakawa, H., Ohara, O. and Owaribe, K.
Novel alternative splicings of BPAG1 (bullous pemphigoid antigen 1) including the domain structure closely related to MacP (microtubule actin cross-linking factor)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLES
JOURNAL
FEATURES
source

J. Biol. Chem. 277 (8), 6682-6687 (2002)
21839111
11751855
2 (bases 1 to 16282)
Okumura, M., Yamakawa, H. and Owaribe, K.
Direct Submission
Submitted (18-JUL-2001) Grad. Human Informatics, Nagoya University,
Unit of Biosystems, Nagoya 464-8601, Japan
Location/Qualifiers
1. 16282
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/cell_type="keratinocyte"
1. 16282
/gene="BPAG1"
1. 15516
/note="BPAG1a; MacP-like product; alternatively spliced"
/product="bullous pemphigoid antigen 1 ea"
/protein_id="AA162061.1"
/db_xref="GI:18157651"
/translation="MHSSSYSSRSDSVFNTSTRTSLDSNENLLVHCPTLINSCLIFGSGSPGHRLEMLQIYANRQVDSDYCEKLLIAGNALQSDSKLESQVQFOEAELIAGYLEBCEENLRHVIDVOILIDGKQVLAQVORVAKLREDEIMALEGSSVSKGILTTBOTKMLISGILTOSINSGSPQTLAPSLTGLTOSLTPSLTSSMTSGLSGMSLTSLTSVTPATTPGPPSGVLPFPSSGVENSLQTLKQNRPLRLKSLLDQNLTEELNMRVQDLNVDENQVLDRTWGSGLPSVSHLENKKNVRALEBESSLKEAKTSEIOMTAPLKLTYARKLRLESQYAKLNTSSNORSHLDLNFPSRAINELIWLKEEBEEVAYDWSEERTNIARKDYHAELEMLDKEENRHSVQEIAGLELLENHARLTI BAYRAMOTOMSMILQICOCVEOHIKENYAPFPNDAAEDYLENLDAIQRKSCDRSSILHKLEDEIWOSEMEKEKELLYKSTIYALMGAKTILOLKRNSCPKLTSPICAKICVROYELITITTKDDCEVLANNHRAKVIISPGNAMPVSCFPTPPNKEAYDLANRIBQOYONVLTLMHBSHINMKSIVSHVYLINEIDRIRASNVASIKTMLGEEHQVLSNLDSRPFEDLESQSVFSGSDITOLEKVNVCQKYOBLKSAEERLEBQESVYNLYISEVRNIRLENCEDRLIRQITRTPLERDDLHBSVFRITTEOEKLEKLEBKDLQDGTINKCEEFPSQAAASSVPTLRSELNVVLQNMNOVSMGSTYIDKLKTVNLKNTQAAALVYKLEETLCREBAVYADKNNIENLISLTKMRSVDKRRQFHALDELOKAKALISDEMTKYRKRDLPFMHKEKAQOLVERNONYVQIDNRLRDLDEIGIKSLKYRTDHYPLDWMIOQVETTRKTIQENPENSKTALATQNLQCKMLVSEIEMKOSMDEOCRTAESQYATVEDELELQNTMYRAMVDSQOKSPVRRMRMSADLIQEFMDLRRYALVTLMTQYIKFAPGAQHEKTIISOHQAVALIATQSAQVLELQGOYLSPBEKELQKMKLEKHVETKOKIISSEESTKQDSEALOTIOFLAGHGMPTDEBNELEKQVTLQESYNLFSBSLKOLOSQTSQDVYKVEKTIYABROQEKELQGIQDILTQTERNLIGHOAPMIGDGTVELKTKOSKQEBELQKMQSQAQALAEVYKTNFPLENKGKLSQEBQALIEQKLNBAKICEQMLKAEQSKLELDKVVTTAKIETKVAVAQVLEBSKTIENLMLNSVQDKOSERAGTKHQVIEQNGTHPQEGGKSAIGEDVNGMLLTDVQGVGTQENLNOQYQKVAQHEKTIISOHQAVALIATQSAQVLELQGOYLSPBEKELQKMKLEKHVETALAESBKMKLTHSLOEBLEKFDADYTEREHMLQOSELEENLEAGADINGLMELKLRKRSRSEDIYSHKGLRYTTISGNRVLEAKSCSRDGGKQVTSATHPVOKLQDHATDPFRSLKYKCVLGNMKLDVNDKYOHYEDASCGLIAGLOACATSKSLSEPIANDPKNIQROULETKALQEQIISQOQVAVELKTAELVLARQSLAPKANDIOCTLTDIADRYEDLSKSVNERNEKLIQITLRSLSVDQDEMDNMGNVSSLJKEGOVPLNSTALQDIISKINMLEQDILAGROSSINAMNEKYKEMETTPSTASLQKMDLSARFSEASHIKETTLAKMEELKTVLELLENLSEKLOTPLETKQALDEVDPKQVTELSQVQSGESPLEHKHLELVLSLKEIISSHGLSPDKLVLEKTNALSKKPEMEDTJKEKEVATSCQEQDLARQVLYVSLSKSIKETTKYPIYQPSFGAEDLQKSLSEDTKLOERKSLKPEIKQVNNSGISLGNLISAVTTPAVALIAVSSGVALVEGEATVTEFPMANKGLTSLIKQDMTIDSHGYEDLGLKDKTALNTLTKLQKQAESSAMQMLQNMKNTLQVTLMOQPAPTDAVTVQVQONKSPFALBKONVKKLMDKLTLELEBNPDPVAPRMQMLTEIDSKQEBLNQLTIDROOKLEESNNLTQFQVAVQKLMVBEKLVAVSVGLPSIDNMLNTQROVQOILLOEFATRRPOVEOLITLAGOGLISLRPGDPSLRGIVKQDLAVTQMDSLITGQLSDRCMTIDALIVKSTQYQSLASLSKLSLDRDKLSSSLAVSTHPRANQOLETAQKQKDEIQDEKQIYKVAOLCELDLALVEEVLKELSLQSLGSLSPQVDEKAEHNHVLQSAACASHQFOOMSRDPQAMVLPDKKEBQONKSPIDKADVLESLLKDNHDFSKTLLAQSAMYEKTIAGESNLTKTGSEKALQOQNLNTIKNMPTFNPQOVERENKLKESLKLKXKQEVETLMPWIDCONNLEIEFCDDPAEGESSIKKLSLOEMQDFGMVELLNTANSILSVCEIDKGVYTPDKSKLOXQYDMTEQOLSHSKFCELENNQTKEFQEVESKSRQLOQCKEODLHDSIGQYSKTYITMLQDTQOKSLQALGQVQDLARLACQDLVAVESDSKSTGVLLQVETLIAQHSSTIQVDEKSKFLETKLQGIQGHPRNTIREMSPFAEPFDELDMSMAVGRDAETLQKQETIKAFLEKLKLEALMASNDNANKTCQMLA

[illegible]

ORIGIN

Query Match	81.3%	Score 1652;	DB 9;	Length 16282;
Best Local Similarly	95.1%	Pred. No. 0;		
Matches 1802; Conservative	0;	Mismatches 75;	Indels 18;	Gaps 9;

Qy	8	GCTGACAGAAACAAATTCATCTCTGACG-TCGCGCCGACCCCTTGAAATATGATAC	66
Db	10620	GCTACAGAAACAAATTCATCTCTGACCTTCCGCGCCACCCCTTGAAATATGAAAC	10667
Qy	67	TCTAAGCAGCAGCAGGAAGAAACAATCGGCACTGCGTATGATGTAGCTGAACAAGCC	126
Db	10680	TCTAAGCAGCAGCAGGAAGAAACAATCGGCACTGCGTATGATGTAGCTGAACAAGCC	10733
Qy	127	TCATATAGATTAAGATGAACAAACTGG----CCAAGTACTGAATGAGCCCTGGGGAGG	182
Db	10740	TCATATAGATTAAGATGAACAAACTGGGCCACAGTTACTGGAATTGAGCCCTGGGGAGG	1079
Qy	183	CTTTTCTATCCAAGAAAGTATGTGGCAGCCGACACCTTTACGTCAATTAAAGAAGA	242
Db	10800	CTTTTCTATCCAAGAAAGTATGTGGCAGCCGACACCTTTACGTCAATTAAAGAAGA	1085
Qy	243	TGTCAAAAGGCTGTGGCAGCTGGAAGGCAATTTCTCAATCACTCACTGCTCCATGA	302
Db	10860	TGTCAAAAGGCTGTGGCAGCTGGAAGGCAATTTCTCAATCACTCACTGCTCCATGA	1091
Qy	303	CAGATAGATCAGATCCTTGAGAGCCTGGAAGCATGCTGAGAGCAGCCACC	362
Db	10920	CAGATAGATCAGATCCTTGAGAGCCTGGAAGCATGCTGAGAGCAGCCACC	1097
Qy	363	CTCATCTCTGACAGGTTGAGGAATCAAGAAACAGATCAGTGAATAATGAATGTGTC	422

Db	10980	CTCTATCTCTGCAGAGCTTGAGAAATCAAGAACAGATCACTGAAAAATTAGATGTGTC	11039
Oy	423	AGTAGACATGGAAAAAGCTACAGCCGTTGTATGAAACTCTTAAACAGAGGGAGAGAAAT	482
Db	11040	AGTAGACATGGAAAAAGCTACAGCCGTTGTATGAAACTCTTAAACAGAGGGAGAGAAAT	11099
Oy	483	GATTGCTTGAATCTGGGGGGGACTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT	542
Db	11100	GATTGCTTGAATCTGGGGGGGACTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT	11159
Oy	543	TGACCAAAATGGTTTTTCATTTGGGGGAACATATCAACACTGGGGAAGAGAGGAAAGCCAA	602
Db	11160	TGACCAAAATGGTTTTTCATTTGGGGGAACATATCAACACTGGGGAAGAGAGGAAAGCCAA	11219
Oy	603	ACTACTGGATGTGATGAGCTAGCAGAAAGATTCTGTGTGATCATATGCTCATTTATAGT	662
Db	11220	ACTACTGGATGTGATGAGCTAGCAGAAAGATTCTGTGTGATCATATGCTCATTTATAGT	11279
Oy	663	TACCATTTAAGATTACTCAAGATTTTCATCCGGGACCTGGAGATCCTGGAAATTAATCTTTC	722
Db	11280	TACCATTTAAGATTACTCAAGATTTTCATCCGGGACCTGGAGATCCTGGAAATTAATCTTTC	11339
Oy	723	AGTAGTAAACAAACAGCAGAAAGCAGACGACCACTAAGGGAGAAATAGATGACTCA	782
Db	11340	AGTAGTAAACAAACAGCAGAAAGCAGACGACCACTAAGGGAGAAATAGATGACTCA	11399
Oy	783	GGAGAGCTGGATTAAGTATTAACTTAGTTCCTGAATCTCAATTTGGGCGATGTGGGAGCC	842
Db	11400	GGAGAGCTGGATTAAGTATTAACTTAGTTCCTGAATCTCAATTTGGGCGATGTGGGAGCC	11459
Oy	843	TGATTAACCCATTGTCAAGAAAGATTAATCATGATTAATTCAGGATGGGATTTCTTAA	902
Db	11460	TGATTAACCCATTGTCAAGAAAGATTAATCATGATTAATTCAGGATGGGATTTCTTAA	11518
Oy	903	ATAAAGCTTGGAAAGACCGGATTGACCACTTGAG-----AGCAATGCACTGCCGTC	956
Db	11519	ATAAAGCTTGGAAAGACCGGATTGACCACTTGAGAGGAGCAATGCACTGCCGTTCACT	11578
Oy	957	GTACAGATGAGCTCAGGCGGTATTTTTCTGGGTGATTAATTCAGAGTGGTAACTTCGCTT	1016
Db	11579	ACCAGGATGAGCTCAGGCGGTATTTTTCTGGGTGATTAATTCAGAGTGGTAACTTCGCTT	11638
Oy	1017	CAATGTCTCCAAATTGGAACAGCATCTCGMAACTGTCAAGCAGCAGATTGATGTGCTAGAG	1076
Db	11639	CAATGTCTCCAAATTGGAACAGCATCTCGMAACTGTCAAGCAGCAGATTGATGTGCTAGAG	11697
Oy	1077	CAATTTAAGTCTGAGGCTTATCAACAGCAGATAG-AGTGTATCGACTATCTATCAACGCA	1135
Db	11698	CAATTTAAGTCTGAGGCTTATCAACAGCAGATAGAAATGGAAGACTAATCATCAAGCA	11757
Oy	1136	GACCCCTTTGCTAAGAAAGTAAACAGAAAGAGAGCAAAACACTGTTCTTAGACCCAAA	1195
Db	11758	GA-CCCTTTGCTAAGAAAGTAAACAGAAAGAGAGTCAAAACACTGTTCTTAGACCCAAAT	11816
Oy	1196	AATGAGACTGAAATTTGATAT-GGCTAGGCTGAGAGAGAAATCATCAACAGACACATTA	1254
Db	11817	AATGAGACTGAAATTTGATATGTGGATAGGCTGAGAGAGAAATCATCAACAGACACATTA	11876
Oy	1255	ACTGAGGGGTCTCTAATTAGCCMTGGGTCAAGTTCACACATGCCCCCTGTATGAGTCTTG--	1312
Db	11877	ACTGAGGGGTCTCTAATTAGCCMTGGGTCAAGTTCACACATGCCCCCTGTATGAGTCTTGCC	11936
Oy	1313	CATGCTGACACACACCGAGGGCTTGCTAAAGTGAGAGAAACCTGTTTGAAGAGAACCTTAA	1372
Db	11937	ATGCTGACACACACCGAGGGCTTGCTAAAGTGAGAGAAACCTGTTTGAAGAGAACCTTAA	11996
Oy	1373	AGCATTGAAATTTAACTTTGCCAAGCATCATGTGCTCAAAATGATGTATTAGCCATCA	1432
Db	11997	AGCATTGAAATTTAACTTTGCCAAGCATCATGTGCTCCAAAATGATGTATTAGCCATCA	12056
Oy	1433	GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTGTAATCAAGTCAGAGCA	1492
Db	12057	GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTAATTGTAATCAAGTCAGAGCA	12116

1111
GAGA 12116

```
QY      1493 AGAAGCAAGCACTTCAGAACAGCTAGAGTTTAAATCAAAGCTGCGAAATGTTTT 1552
      12117 AGAAGCAAGCACTTCAGAACAGCTAGAGTTTAAATCAAAGCTGCGAAATGTTTT 12176
QY      1553 GGAAGAAAACAGAACAAAGAGAGAGAGCTGATGCTGCTTGCGCCAAAGGCTT 1612
      12177 GGAAGAAAACAGAACAAAGAGAGAGAGCTGATGCTGCTTGCGCCAAAGGCTT 12236
QY      1613 CCATGGCGAAATTGAGGATTTGAGAGCTGAGCTGACGAGAGCTGATCTGTTGCG 1672
      12237 CCATGGCGAAATTGAGGATTTGAGAGCTGAGCTGACGAGAGCTGATCTGTTGCG 12296
QY      1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGAGAGCTTATGTCATATGA 1732
      12297 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGAGAGCTTATGTCATATGA 12356
QY      1733 AGTCTGCTGCTCTTTGAGCTTAAAGAAACATATAGAGTCTGATGAGAAAGCCA 1792
      12357 AGTCTGCTGCTCTTTGAGCTTAAAGAAACATATAGAGTCTGATGAGAAAGCCA 12416
QY      1793 GCAGATGCTTGCAAGATGCCCAAAATCTGAGAGACAAATATGACCAAGACATTAATA 1852
      12417 GCAGATGCTTGCAAGATGCCCAAAATCTGAGAGACAAATATGACCAAGACATTAATA 12476
QY      1853 CTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAC 1887
      12477 CTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAC 12511
Db      12477 CTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAC 12511

RESULT 4
AX834329      3307 bp      DNA      linear      PAT 15-DEC-2003
LOCUS      AX834329
DEFINITION      Sequence 1453 from Patent EP1347046.
ACCESSION      AX834329
VERSION      AX834329.1 GI:39920464
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamehika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuko,Y.
TITLE      Full-length cDNA sequences
JOURNAL      Patent: EP 1347046-A 1453 24-SEP-2003;
            Research Association for Biotechnology (JP)
FEATURES
            source
            1..3307
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      80.6%; Score 1637.6; DB 6; Length 3307;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

QY      8 GCTGACAGAAACAATCAATCATCTCTCAGC-TCCGCGCCAGCCCTTGATATGATAC 66
      1457 GCTGACAGAAACAATCAATCATCTCTCAGCTTCCGCGCCAGCCCTTGATATGATAC 1516
Db      1457 GCTGACAGAAACAATCAATCATCTCTCAGCTTCCGCGCCAGCCCTTGATATGATAC 1516
QY      67 TCTAAGGCGACGACAGAGAGACATGCGCACTGCGTGAATTGATGACGACCAAGCC 126
      1517 TCTAAGGCGACGACAGAGAGACATGCGCACTGCGTGAATTGATGACGACCAAGCC 1576
Db      1517 TCTAAGGCGACGACAGAGAGACATGCGCACTGCGTGAATTGATGACGACCAAGCC 1576
QY      127 TCATATAGATTAATGAAACAACTGG-----CCACAGTACTGAAATGAGCCCTGGGAGG 182
      1577 TCATATAGATTAATGAAACAACTGGGCGACAGTTACTGAAATTTGAGCCCTGGGAGG 1636
Db      1577 TCATATAGATTAATGAAACAACTGGGCGACAGTTACTGAAATTTGAGCCCTGGGAGG 1636
QY      183 CTTTCTATCCAAGAGAGATGATGCGACCGACACCTTTACAGTCACTTAAGAGAG 242
```

```
Db      1637 CTTTCTATCCAAGAGAGATGATGCGACCGACACCTTTACAGTCACTTAAGAGAG 1696
      243 TGTCAAAAAGCGTGTGGGCACTGATGAGCGATTTCTCAATCAACTCAGTTCATGA 302
QY      1697 TGTCAAAAAGCGTGTGGGCACTGATGAGCGATTTCTCAATCAACTCAGTTCATGA 1756
Db      303 CAAGATAGATCAGATCTTGAAGCGCTGGAAAGCATGCGTGAACGTCGAGGACGACC 362
      1757 CAAGATAGATCAGATCTTGAAGCGCTGGAAAGCATGCGTGAACGTCGAGGACGACC 1816
QY      363 CTCATCTCTGCAAGGTTGAAGATCAAGAAACGATCAGTGAATTAAGATGTGC 422
      1817 CTCATCTCTGCAAGGTTGAAGATCAAGAAACGATCAGTGAATTAAGATGTGC 1876
Db      423 AGTAGCAATGAAAAAGCTACAGCCCTGTATGAAACTCTTAAACAGAGGGAGAGAAAT 482
      1877 AGTAGCAATGAAAAAGCTACAGCCCTGTATGAAACTCTTAAACAGAGGGAGAGAAAT 1936
QY      483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAGCT 542
      1937 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAGCT 1996
Db      543 TGAACCAATGTTTTTCAATTTGGGAGAACATACACACTGTGGAAGAGAGGAGCCAA 602
      1997 TGAACCAATGTTTTTCAATTTGGGAGAACATACACACTGTGGAAGAGAGGAGCCAA 2056
QY      603 ACTACTGATGATGATGAGTACTGACGAGAAAGTTCTGCTGTGATCAATGTCAATGATG 662
      2057 ACTACTGATGATGATGAGTACTGACGAGAAAGTTCTGCTGTGATCAATGTCAATGATG 2116
Db      663 TACCATTAAGATTAATCAAGATTTTCACTCCGGACCTGGAAGATCTCGAATTTGATCTTC 722
      2117 TACCATTAAGATTAATCAAGATTTTCACTCCGGACCTGGAAGATCTCGAATTTGATCTTC 2176
QY      723 AGTAGTAACCAACAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
      2177 AGTAGTAACCAACAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2236
Db      783 GAGAGAGCTGATATATGATTTATTAATCACTAGTCTGAACTCATTTGCCGACTGTGGGAGCC 842
      2237 GAGAGAGCTGATATATGATTTATTAATCACTAGTCTGAACTCATTTGCCGACTGTGGGAGCC 2296
QY      843 TGATTAACCCATTTGTAAAGAGATATCATGATGTTAAATTCAGC-ATGGGATTTCTCTAA 902
      2297 TGATTAACCCATTTGTAAAGAGATATCATGATGTTAAATTCAGC-ATGGGATTTCTCTAA 2355
Db      903 ATTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 956
      2356 ATTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2415
QY      957 GTACAGATGAGATGCGAGGCGGATTTTCTGGGTAGATTTTGCAGGTGTGTAAGTTGCTT 1016
      2416 ACCAGATGAGATGCGAGGCGGATTTTCTGGGTAGATTTTGCAGGTGTGTAAGTTGCTT 2475
Db      1017 CAATGCTCCAAATTTGGAACAGAGATCTGAAACGTCGAACAGAGAGATGATGCTAGAG 1076
      2476 CAATGCTCCAAATTTGGAACAG-ATCTCGAAACTGTCAACAGAGAGATGATGAGCTAAG 2534
QY      1077 CAATTTAAGCTGAGGCTTATCAACAGCAGATAG-AGTGTAAGCTGACGTCATCAACGA 1135
      2535 CAATTTAAGCTGAGGCTTATCAACAGCAGATAGAAATGGAAGAGCTGAATCATCAAGCA 2594
Db      1136 GACGCTTTTGTAAAGAAAGTAACAGAGAGAGTGAACAAACACTGTTCTAGACCCAA 1195
      2595 GA-GCTTTTGTAAAGAAAGTAACAGAGAGAGTGAACAAACACTGTTCTAGACCCAA 2653
QY      1196 AATGAGCTGATTTGATAT-GGCTAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1254
      2654 AATGAGCTGATTTGATATGAGATAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2713
Db      1255 ACTGAGGAGTGTCTATTAGCCNTGGGTGAGTCCAACTAGCTGTATAGCTCTCTG-- 1312
```

Db	2714	ACTGAGGGTGTCTATTAGGCTTGGTCAAGTCCATCCGCTGATGAGTCTCTGGC	2773
Qy	1313	CATGCTGACACACCGAGGGCTTGTAAGAGAGAAACCTGTGGAGAGAGACCCCTAA	1372
Db	2774	ATGCTGACACACACCGAGGGCTTGTAAGAGAGAAACCTGTGGAGAGAGACCCCTAA	2833
Qy	1373	AGCCATTGAATTTGAATCTTGGCAAGCATCATGTGCTCCAAATATGATTTAGCCCATCA	1432
Db	2834	AGCCATTGAATTTGAATCTTGGCAAGCATCATGTGCTCCAAATATGATTTAGCCCATCA	2893
Qy	1433	GTCCACAGTGAAGCGCTTAAATTAAGCAGAAATGATCTAATTAATCACTGACAGAGA	1492
Db	2894	GTCCACAGTGAAGCGCTTAAATTAAGCAGAAATGATCTAATTAATCACTGACAGAGA	2953
Qy	1493	AGAAGCAAGCAACCTTTCAGAACCACTGAGATTTTAAATCAAGCTTGGCAAAATGTTT	1552
Db	2954	AGAAGCAAGCAACCTTTCAGAACCACTGAGATTTTAAATCAAGCTTGGCAAAATGTTT	3013
Qy	1553	GGAAGAAAACAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1612
Db	3014	GGAAGAAAACAGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3073
Qy	1613	CCATGGCGAAATTGAGATTTCAGACAGTGGCTGACTGACACGAGCGTCACTCTGGC	1672
Db	3074	CCATGGCGAAATTGAGATTTCAGACAGTGGCTGACTGACACGAGCGTCACTCTGGC	3133
Qy	1673	ATCTAAACCGCTGGAGAGTTTACCGGAAACAGCCAGAGAGAGAGAGAGAGAGAGAG	1732
Db	3134	ATCTAAACCGCTGGAGAGTTTACCGGAAACAGCCAGAGAGAGAGAGAGAGAGAGAG	3193
Qy	1733	AGTGTGTGCTGCTTTTGAAGCTTAAGAAAGAAATATATAGAGTCTGATGACAGAAAGGCA	1792
Db	3194	AGTGTGTGCTGCTTTTGAAGCTTAAGAAAGAAATATATAGAGTCTGATGACAGAAAGGCA	3253
Qy	1793	GCAGATCTTTCAGAGATGCCCAAAATTCGACAGAGACAAATATTCACCAAGACAT	1846
Db	3254	GCAGATCTTTCAGAGATGCCCAAAATTCGACAGAGACAAATATTCACCAAGACAT	3307

RESULT 5
 AK096713 3307 bp mRNA linear PRI 30-JAN-2004
 LOCUS Homo sapiens cDNA FLJ39394 f1s, clone PLACE6007925, moderately
 DEFINITION similar to Homo sapiens mRNA for actin binding protein ABP620.
 AK096713
 VERSION AK096713.1 GI:21756267
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, H.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
 Shiratori, A., Sudo, H., Hosioki, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, E., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuno, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
 Murakawa, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M.,
 Hirooka, S., Chiba, Y., Ishida, S., Oho, Y., Takiguchi, S., Watanabe, S.,
 Yoshida, M., Hottu, T., Kusano, J., Kanehori, K., Takahashi, F., Hara, R.,
 Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Ohnima, A.,
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunaga, H., Ichihara, T.,
 Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satch, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
 Teshit, H., Tanigami, A., Fujitawa, T., Ono, T., Yamada, K., Fujii, Y.,

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
 Inagaki, S., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, K.,
 Noguchi, S., Itoh, T., Shigeta, K., Senda, T., Matsunaga, T.,
 Nakajima, Y., Mizuno, T., Norinaga, M., Sasaki, M., Togaishi, T.,
 Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
 Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
 Nagase, T., Nomura, N., Kikkawa, K., Masuo, Y., Yamashita, R.,
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 14702039

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y.,
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Ohnima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahara, K., Masuo, Y., Nagai, K. and Isogai, T.
 NEBO human cDNA sequencing project
 3 (bases 1 to 3307)
 Unpublished
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazuo-Kamataxi, Katsarazu, Chiba 292-0612, Japan
 (E-mail: genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEBO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology (RAB); cDNA library
 construction; Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' & 3' end one pass sequencing; RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing; HRI and
 RAB; annotation; HRI and RAB.

FEATURES
 source
 Location/Qualifiers
 1..3307
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE6007925"
 /cissue="placenta"
 /clone_lib="PLACE6"
 /note="cloning vector: pMB18SFL3"
 336..3307
 /note="unlabeled protein product"
 /codon_start=1
 /protein_id="BAC04848.1"
 /db_xref="GI:21756268"
 /translation="MVSQKDFQVWYIEIOEKSHSRLDQALCNKIFGEDEVET
 MNINLVHDKSLSTVODSTEGMLQSESLRLOEIIILRNQNDALNGLELQ
 TTDDEVLTIDKLEIKARVYKDTSTVDAKLBOLALCRLSHETSLRMDV
 EVLLSTVETVLKGBEASQAMPKPKSKAKNNKALDLSVSSALLELVMPRRS
 GLEKVAEDERLRLVSDITQVVEIDAILLSQFPDQADNLSLITETKLSL
 GDRLRQDQVADLOVQKTFMTLILHKKIDIDLVSGHMTKACSEERQSKALYD
 KVLNVDTCQINSEVRLQERASLVNCPWETBELMPULTEVQSIISLPALRY
 IETROOESHRLOREILAEHPHIDKNNKGPOLLESPESGSEIOKRYAATPLVSQ
 IKEDVRAVALDEALISQHPHIDKNNKGPOLLESPESGSEIOKRYAATPLVSQ
 ENQVSVNDEKLOPLVETLKQREEMTASGGTDDKISAKAVDDKQDQNTENIT
 LVEERKMLDVELLEKFCWCDMSLITVTKQDTRIDEDGIDSPSVKQOQEA
 TIBERIDGLOEELDIVINSEELIACGEBDKIVKSIIDELNASDLSNKKQD
 KLEBAQAQVQDQVADVNDVIGAKLSASPIGTDETVKQOIEELKQSEKAY
 QOIEEMRLNHOEELKKVTESSDGHVODPLMEKLIWDSLEERTINRQKLEGL
 LAQGOHNLDELALVLTREGLTSKQPGDPCDPAIEIELAGHYLQNDVLAHOSV
 EAVKRGNDLISSEASBASNQKLEVLNQRQNYLKTGRQKQDGLDQKRAKGRH
 GRVEDIQWITDTERHLASKPLGIGIPETAKQOLNVHVECAFAEAKETYSKLMQK
 QQMLACPKSAETNIDQD"

ORIGIN
 Query Match 80.64; Score 1637.6; DB 9; Length 3307;

Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

```

QY      8 GGTGACGAAACCAATCAATCATCTCTCAGC--TCCGCCCCAGCCCTTGAATATGATAC 66
Db     1457 GCTGACGAAACCAATCAATCATCTCTCAGCTTCCGCCCCAGCCCTTGAATATGAAAC 1516
QY      67 TCTAAGGACGACGAGGAAAGACATCGGCACTCGGTGAGTTGATAGCTGAAACAAGACC 126
Db     1517 TCTAAGGACGACGAGGAAAGACATCGGCACTCGGTGAGTTGATAGCTGAAACAAGACC 1576
QY     127 TCAATATGATTAAGTGAACAAAACCTGG---CCAAGTACTGATATGAGCCCTGGGGAAG 182
Db     1577 TCAATATGATTAAGTGAACAAAACCTGGGCCACAGTTACTGGAATTGAGCCCTGGGGAAG 1636
QY     183 CTTTTCATCAAGAGAAAGTATGTGGCAGCCGACACCCCTTACAGTACACTTAAAGAGA 242
Db     1637 CTTTTCATCAAGAGAAAGTATGTGGCAGCCGACACCCCTTACAGTACACTTAAAGAGA 1696
QY     243 TGTCAAAAAGCGTGTGTGCACTGATGAAAGCCATTTCTCAATCAACTCAGTTCCATGA 302
Db     1697 TGTCAAAAAGCGTGTGTGCACTGATGAAAGCCATTTCTCAATCAACTCAGTTCCATGA 1756
QY     303 CAAGATATGATCAATCTCTTGAGAGCTGGAACGCAATCTGGAACGTCTGAGGACCCACC 362
Db     1757 CAAGATATGATCAATCTCTTGAGAGCTGGAACGCAATCTGGAACGTCTGAGGACCCACC 1816
QY     363 CTCTATCTCTGCAAGAGTTGAGAAGATCAAGAAACGATCAGTGAATAATGAATGTGTC 422
Db     1817 CTCTATCTCTGCAAGAGTTGAGAAGATCAAGAAACGATCAGTGAATAATGAATGTGTC 1876
QY     423 AGTAGACATGAAAAAGCTACAGCCGTTGTATGAAACTCTTAAACAGAGGGAGAGGAAT 482
Db     1877 AGTAGACATGAAAAAGCTACAGCCGTTGTATGAAACTCTTAAACAGAGGGAGAGGAAT 1936
QY     483 GATTGCTAGATCTGGGGGGGACTGATTAAGCATATCTGCCAAAGCTGTTCAAGATTAAGT 542
Db     1937 GATTGCTAGATCTGGGGGGGACTGATTAAGCATATCTGCCAAAGCTGTTCAAGATTAAGT 1996
QY     543 TGACCAAAATGCTTTTCACTTTGGGAGAACATACACACCTGTGGAAAGAGGGAGAGCCAA 602
Db     1997 TGACCAAAATGCTTTTCACTTTGGGAGAACATACACACCTGTGTGAAGAGGGAGAGCCAA 2056
QY     603 ACTACTGATGTATGAGAGCTAGACAGAAAAATCTGTGTGATCACAATGTCTATGATAGT 662
Db     2057 ACTACTGATGTATGAGAGCTAGACAGAAAAATCTGTGTGATCACAATGTCTATGATAGT 2116
QY     663 TACCATTTAAAGATTAATCTCAAGATTTTCAATCCGGGACCTGGAAGATCTGGAAATTGATCTTC 722
Db     2117 TACCATTTAAAGATTAATCTCAAGATTTTCAATCCGGGACCTGGAAGATCTGGAAATTGATCTTC 2176
QY     723 AGTAGTAAAAACAACAGCAAGACAGACAGACCATTAAGGGAAAGAAATTAATGATGATCA 782
Db     2177 AGTAGTAAAAACAACAGCAAGACAGACAGACCATTAAGGGAAAGAAATTAATGATGATCA 2236
QY     783 GGAAGAGCTGAGTATATGTTATTAACCTAGGTTCTGAACTGATTCGGGCAATGTGGGAGCC 842
Db     2237 GGAAGAGCTGAGTATATGTTATTAACCTAGGTTCTGAACTGATTCGGGCAATGTGGGAGCC 2296
QY     843 TGTATAAACCATTTGTCAAGAGAGTATCAATGATTAATTTCAAGCGATGGGATTTCTTAA 902
Db     2297 TGTATAAACCATTTGTCAAGAGAGTATGATGATTAATTTCAAGC-ATGGGATTTCTTAA 2355
QY     903 ATAAAGCTTGGAAAGACCGGATTTGACCAACTTGAG-----AACATATGACGCTGCCGTC 956
Db     2356 ATAAAGCTTGGAAAGACCGGATTTGACCAACTTTAGAGAGCAATGACGCGCTGCCGTTCACT 2415
QY     957 GTACAGATGAGCTGCAAGCGCGTATTTTCTGGGTAGATATTTGACAGGTGTAATTCGCTT 1016
Db     2416 ACCAGATGAGCTGCAAGCGCGTATTTTGAAGGTGATGATTTGACAGGTGTAATTTGAGCTT 2475
QY    1017 CAATGTCTCCAATTGGAAACGCAATCTCGAAACTGTCAAGACGACGATGATGTGCTAGAG 1076

```

```

Db     2476 CAATGTCTCCAATTGGAAACG-ATCTCGAAACTGTCAAGCAGACGATTTGAGACTTAAG 2534
QY    1077 CAATTTAAGTCTGAGGCTTATCAACAGACAGATAG-AGTGGTACGATGACTCATCGCA 1135
Db     2535 CAATTTAAGTCTGAGGCTTATCAACAGCAGATATGAAATGAAAGACTGATATCTCAAGCA 2594
QY    1136 GACGCTTTTGTCAAAGAAAGTAAACAGMAGAGTGAACAACACTGTTCTAGACCCAA 1195
Db     2595 GA-GCTTTTGTCAAAGAAAGTAAACAGMAGAGTGAACAACACTGTTCAAGACCATT 2653
QY    1196 AATGAGCTGATTTTATAT-GGCTAGCTGCGAGAGAGATATATCAACAGACAGATTA 1254
Db     2654 AATGAGCTGAAATTTATATATGAGTATGCTTGAAGAGAAATCATTAACAGACAGATTA 2713
QY    1255 ACTGAGGGGTGCTTATTAAGCCATGAGGTGAGTTCCAACATGACCCTGATAGGCTCTG-- 1312
Db     2714 ACTGAGGGGTGCTTATTAAGCCATGAGGTGAGTTCCAACATGACCCTGATAGGCTCTG-- 2773
QY    1313 CATGCTGACACACACGAGGGCTTGTCTAAGTAGAGAGAAACCTGTTGAGAGAACCTTAA 1372
Db     2774 ATGGCTGACACACACGAGGGCTTGTCTAAGTAGAGAGAAACCTGTTGAGAGAACCTTAA 2833
QY    1373 AGCCATTGAATTTGAACTTGGCAAGCATATGTCTCCAAATATGATATTTAGCCATCA 1432
Db     2834 AGCCATTGAATTTGAACTTGGCAAGCATATGTCTCCAAATATGATATTTAGCCATCA 2893
QY    1433 GTCCACAGTGAAGCCGCTTATTAAGCAGAAATGATCTTAATTTGAATCAAGTGCAGAGAG 1492
Db     2894 GTCCACAGTGAAGCCGCTTATTAAGCAGAAATGATCTTAATTTGAATCAAGTGCAGAGAG 2953
QY    1493 AGAAGCAACCACTTCAAGAACAGCTAGAGTTTAAATCAACGCTGCAAAATGTTT 1552
Db     2954 AGAAGCAACCACTTCAAGAACAGCTAGAGTTTAAATCAACGCTGCAAAATGTTT 3013
QY    1553 GGAAGAAACAGAACAAAGAGACAGACTGATGTGCTTGGCCAGACCCAAAGGTT 1612
Db     3014 GGAAGAAACAGAACAAAGAGACAGACTGATGTGCTTGGCCAGACCCAAAGGTT 3073
QY    1613 CCATGGCGAAATTTGAGAGTTTGGCAGAGTGGGCTGACAGGAGCGTCAATGTTGGC 1672
Db     3074 CCATGGCGAAATTTGAGAGTTTGGCAGAGTGGGCTGACAGGAGCGTCAATGTTGGC 3133
QY    1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAAGAGACAGCTTAATGTCCATATGGA 1732
Db     3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAAGAGACAGCTTAATGTCCATATGGA 3193
QY    1733 AGTCTGTGCTGCTTTGAAAGCTTAAGAAAGAAATTAAGTCTATGACGAAGGCCA 1792
Db     3194 AGTCTGTGCTGCTTTGAAAGCTTAAGAAAGAAATTAAGTCTATGACGAAGGCCA 3253
QY    1793 GCAGATGCTTGCAGAGATGCCCCAAAATCTGCAGAGCAAAATTTGACCAAGACAT 1846
Db     3254 GCAGATGCTTGCAGAGATGCCCCAAAATCTGCAGAGCAAAATTTGACCAAGACAT 3307

RESULT 6
AK128632 4112 bp mRNA linear PRI 19-FRB-2004
LOCUS AK128632
DEFINITION Homo sapiens cDNA FLJ46791 fis, clone TRACH3029462, moderately
similar to Actin cross-linking family protein 7.
ACCESSION AK128632
VERSION AK128632.1 GI:34536103
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kanohori, K., Ishibaishi, T., Chiba, Y., Fujimoto, K., Hirooka, S.,
Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Horiuchi, T., Watanabe, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Oka, T., Wakamatsu, A.,
Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hiro, Y., Saito, K.,
Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M.,

```


Kikuchi, H., Kande, K., Watsushima, M., Takahashi-Fujii, A., Oshima, A.,
 Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahara, K.,
 Masuho, Y., Nagai, K. and Iseogai, T.
 NEDO human cDNA sequencing project
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Submitted (15-JUL-2003) Takao Iseogai, FRI Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@kri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology (RAB) (supported by Japan
 Key Technology Center etc.); 5' - 3' - end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: Reverse Proteomics Research Institute, HRI and
 RAB.

FEATURES

Source

Location/Qualifiers
 1..4112
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TRACH3029462"
 /cissue_type="trachea"
 /clone_11b="TRACH3"
 /note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 80.6%; Score 1637.6; DB 9; Length 4112;
 Best Local Similarity 95.9%; Pred. No. 0;

Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

QY 8 GCTGACAGAAACACATCATCTCTCAGC-TCGCCGCCAGCCCTTGAATATGATAC 66
 DB 2262 GCTGACAGAAACACATCATCTCTCAGCCTTCCGCCCTTGAATATGATAC 2321
 QY 67 TCTAAGCAGCAGCAGAGAAACATCGCACTGCGTATGATAGCTGAAACACAGCC 126
 DB 2322 TCTAAGCAGCAGCAGAGAAACATCGCACTGCGTATGATAGCTGAAACACAGCC 2381
 QY 127 TCATTTAGATAGATGAAACAAACCTGC---CCACAGTACTGATAGACCTCGGGAAAG 182
 DB 2382 TCATTTAGATAGATGAAACAAACCTGC---CCACAGTACTGATAGACCTCGGGAAAG 2441
 QY 183 CTTTCTCTCCAGAGATATGTGGCAGCCGACACCTTTACAGTCACTTAAAGAGA 242
 DB 2442 CTTTCTCTCCAGAGATATGTGGCAGCCGACACCTTTACAGTCACTTAAAGAGA 2501
 QY 243 TGTCAAAAAGCGTGTGTGGCAGTGTGAAGCCATTTCTCAATCACTCACTTCAATGA 302
 DB 2502 TGTCAAAAAGCGTGTGTGGCAGTGTGAAGCCATTTCTCAATCACTCACTTCAATGA 2561
 QY 303 CAAATATAGATAGATCTTGTGAAGCGCTTGAAGCGATCTGTGAAGCGCTTGAAGCGCC 362
 DB 2562 CAAATATAGATAGATCTTGTGAAGCGCTTGAAGCGATCTGTGAAGCGCTTGAAGCGCC 2621
 QY 363 CTTATCTCTGCAAGGTTGAAGAGATCAAGAAACAGATCACTGAAATTAAGATATGTGTC 422
 DB 2632 CTTATCTCTGCAAGGTTGAAGAGATCAAGAAACAGATCACTGAAATTAAGATATGTGTC 2681
 QY 423 AGTAGACATGAGAAAGCTACAGCCGTTGTATGAAACCTTTAAACAGAGGGAGAGAAAT 482
 DB 2682 AGTAGACATGAGAAAGCTACAGCCGTTGTATGAAACCTTTAAACAGAGGGAGAGAAAT 2741
 QY 483 GATTGCTATGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAGCT 542
 DB 2742 GATTGCTATGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAGCT 2801
 QY 543 TGACCAATATGCTTTTCATTTGGAGAACATACACACACTGGTGAAGAGAGAGCCAA 602
 DB 543 TGACCAATATGCTTTTCATTTGGAGAACATACACACACTGGTGAAGAGAGAGCCAA 602

DB 2802 TAACCAATATGCTTTTCATTTGGAGAACATACACACACTGGTGAAGAGAGCCAA 2861
 QY 603 ACTACTGATATGATGAGAGCTAGACAGAAAGTTCTGGTGTATCATATGATGATAGT 662
 DB 2862 ACTACTGATATGATGAGAGCTAGACAGAAAGTTCTGGTGTATCATATGATGATAGT 2921
 QY 663 TACCATTAAGATATCTCAAGATTTTCATCCGGGACCTGGAAGATCTGGAAATGATCTCTC 722
 DB 2922 TACCATTAAGATATCTCAAGATTTTCATCCGGGACCTGGAAGATCTGGAAATGATCTCTC 2981
 QY 723 AGTAGTAAACCAACAGCAGAGAGAGACCATTAAGAGAGAAATAGATGAGACTACA 782
 DB 2982 AGTAGTAAACCAACAGCAGAGAGAGACCATTAAGAGAGAAATAGATGAGACTACA 3041
 QY 783 GGAGAGCTGATATGATTTATTAACCTTAGCTTGAACCTCATTTGGCGATGTGGAGACC 842
 DB 3042 GGAGAGCTGATATGATTTATTAACCTTAGCTTGAACCTCATTTGGCGATGTGGAGACC 3101
 QY 843 TGATTAACCCATTTGCAAGAAAGATATACATGATTAATTCAGAGATGGGATCTCTAA 902
 DB 3102 TGATTAACCCATTTGCAAGAAAGATATACATGATTAATTCAGC-ATGGATTTCTTAA 3160
 QY 903 ATAAAGCTTGAAGAACCGGATTTGACCAACTTGGAG-----AGCAATGCACTGCCGTCA 956
 DB 3161 ATAAAGCTTGAAGAACCGGATTTGACCAACTTGAAGAGCCAAATGCAAGCTGCCGTCACT 3220
 QY 957 GTACAGATGACCTGACAGCGGATTTTCTGGGTATGATATTTGCAAGGTGTATGTTCCGTT 1016
 DB 3221 ACCAGATATGACCTGACAGCGGATTTTCTGGGTATGATATTTGCAAGGTGTATGTTCCGTT 3280
 QY 1017 CAATGCTCCCAATTGGAACAGCATCTGAAACCTGCAAGCAGCAGATTTGATGTCCTAGAG 1076
 DB 3281 CAATGCTCCCAATTGGAACAGCATCTGAAACCTGCAAGCAGCAGATTTGATGTCCTAGAG 3339
 QY 1077 CAATTTAAGTCTGAGGCTATCAACAGCAGATAG-AGTGTATGACAGTCACTCATCAGCA 1135
 DB 3340 CAATTTAAGTCTGAGGCTATCAACAGCAGATAG-AGTGTATGACAGTCACTCATCAGCA 3399
 QY 1136 GACGCTTTGCTAAAGAAAGTATACAGAGAGAGTGAACAAACACTGTTCTGAGACCCAA 1195
 DB 3400 GA-GCTTTGCTAAAGAAAGTATACAGAGAGAGTGAACAAACACTGTTCTGAGACCCAT 3458
 QY 1196 AATGAGTGAATATTTGATAT-GGCTAGCCTGCAAGAGAGATTCATCAACAGCAGCATAA 1254
 DB 3459 AATGAGTGAATATTTGATATGAGATATGAGATATGAGAGAGATTCATCAACAGCAGCAT 3518
 QY 1255 ACTGAGGAGTCTCTATTAAGCCTGAGTCAATGCTCCATGATGAGTCTCTG-- 1312
 DB 3519 ACTGAGGAGTCTCTATTAAGCCTGAGTCAATGCTCCATGATGAGTCTCTG-- 3578
 QY 1313 CATGCTGACACACAGGAGGCTTTGATGAGCAGAAACCTGTGGAGAGACCTTAA 1372
 DB 3579 ATGCTGACACACAGGAGGCTTTGATGAGCAGAAACCTGTGGAGAGACCTTAA 3638
 QY 1373 AGCATTAAGATTAAGATCTGCAAGCATATGCTCTCAAAATGATTAATAGCCCATCA 1432
 DB 3639 AGCATTAAGATTAAGATCTGCAAGCATATGCTCTCAAAATGATTAATAGCCCATCA 3698
 QY 1433 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTTAATTAATCAATGACAGAGA 1492
 DB 3699 GTCCACAGTGAAGCCGTTAAATTAAGCAGAAATGATCTTAATTAATCAATGACAGAGA 3758
 QY 1493 AGAAGCAAGCAACCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTT 1552
 DB 3759 AGAAGCAAGCAACCTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGGCAAAATGTTTT 3818
 QY 1553 GGAAGAAACAGAAACAGAGAGAGCAGCTGATGATGCTTGGCCCAAGGCAAAAGGTT 1612
 DB 3819 GGAAGAAACAGAAACAGAGAGAGCAGCTGATGATGCTTGGCCCAAGGCAAAAGGTT 3878
 QY 1613 CCATGAGCAAAATGAGATTTGAGCAGAGTGGCTGACAGCAGAGGCTCATCTGTTGGC 1672
 DB 3879 CCATGAGCAAAATGAGATTTGAGCAGAGTGGCTGACAGCAGAGGCTCATCTGTTGGC 3938

QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGACAGCTTATGTCATATGA 1732
|
|
|
Db 3939 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGACAGCTTATGTCATATGA 3998
|
|
|
QY 1733 AGTCTGCTGCTCTTGAAGCTTAAAGAAACAATATAGAGTCTGATGCAAGAGCCA 1792
|
|
|
Db 3999 AGTCTGCTGCTCTTGAAGCTTAAAGAAACAATATAGAGTCTGATGCAAGAGCCA 4058
|
|
|
QY 1793 GCAGATGCTTGCAAGATGCCAAATCTGCAAGAGCAAAATATGACCAAGACAT 1846
|
|
|
Db 4059 GCAGATGCTTGCAAGATGCCAAATCTGCAAGAGCAAAATATGACCAAGACAT 4112
|
|
|
RESULT 7
LOCUS CQ732923 7194 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 18857 from Patent WO02068579.
ACCESSION CQ732923
VERSION CQ732923.1 GI:42314338
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1
TITLE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.M.
Kits, such as nucleic acid arrays, comprising a majority of
humanecons or transcripts, for detecting expression and other uses
of theof
Patent: WO 02068579-A 18857 06-SEP-2002;
PE Corporation (NY) (US)
JOURNAL Location/Qualifiers
FEATURES
source 1..7194
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 65.2%; Score 1324.6; DB 6; Length 7194;
Best Local Similarity 87.0%; Pred. No. 7.1e-311;
Matches 1645; Conservative 0; Mismatches 54; Indels 191; Gaps 10;
QY 8 GCTGACAGAAACAATCATCATCTCTCAGC-TCCGCGCCAGCCCTTGAATATGATAC 66
|
|
|
Db 2580 GCTGACAGAAACAATCATCATCTCTCAGC-TCCGCGCCAGCCCTTGAATATGATAC 2639
|
|
|
QY 67 TCTAAGCGACGACGAGAGAACATCGGCACTGCGTGAAGTTGATAGCTGAACAACAGCC 126
|
|
|
Db 2640 TCTAAGCGACGACGAGAGAACATCGGCACTGCGTGAAGTTGATAGCTGAACAACAGCC 2699
|
|
|
QY 127 TCATATAGATTAATGATGAACAAACTGG---CCACAGTACTGTAATAGCCCTGGGGAAG 182
|
|
|
Db 2700 TCATATAGATTAATGATGAACAAACTGGGCCACAGTTACTGGAATTAAGCCCTGGGGAAG 2759
|
|
|
QY 183 CTTTCTATCCAAGAGAATATGTGGCAGCGCACACCTTTACAGTCACTTAAGAAGA 242
|
|
|
Db 2760 CTTTCTATCCAAGAGAATATGTGGCAGCGCACACCTTTACAGTCACTTAAGAAGA 2819
|
|
|
QY 243 TGTCAAAAAGCGTGTGTGCACTGATGAAGCCATTTCATCAATCAACTCAATGA 302
|
|
|
Db 2820 TGTCAAAAAGCGTGTGTGCACTGATGAAGCCATTTCATCAACTCAATGA 2868
|
|
|
QY 303 CAAGATATGATCAATCTTGAGAGCTTGAAACGATCTGTGAACGCTGTGAGGACGCCACC 362
|
|
|
Db 2869 ----- 2868
|
|
|
QY 363 CTCTATCTCTGAGAGGTTGAGAAGATCAA-GGAACGATCAGTGAATAATGAATGTG 421
|
|
|
Db 2869 -----CAGAGGTTGAGAAGATCAAAGGAAACGATCAATGAATAATGAATGTG 2917
|
|
|
QY 422 CAGTAGACATGAAGAAAGCTACAGCCGTTGTATGAATCTTTAAACAGAGGGAGAGAGAA 481
|
|
|

Db 2918 CAGTAGACATGGAAGAAAGCTACAGCCGTTGTATGAATCTTTAAACAGAGGGAGAGAGAA 2977
|
|
|
QY 482 TGAATGCTGATCTGGGGGGAAGCTATTAAGACATATCTGCAAGCTGTTACGATTAAG 541
|
|
|
Db 2978 TGAATGCTGATCTGGGGGGAAGCTATTAAGACATATCTGCAAGCTGTTACGATTAAG 3037
|
|
|
QY 542 TTGACCAAAATGTTTTTCATTTGGGAGAACATACACACACTGTGTGAAGAGAGGAAAGCCA 601
|
|
|
Db 3038 TTGACCAAAATGTTTTTCATTTGGGAGAACATACACACACTGTGTGAAGAGAGGAAAGCCA 3097
|
|
|
QY 602 AACTACTGATGTGATGAGCTTACGAGAAAGTTCTGTGTGTATCACTGTCATTTATAG 661
|
|
|
Db 3098 AACTACTGATGTGATGAGCTTACGAGAAAGTTCTGTGTGTATCACTGTCATTTATAG 3157
|
|
|
QY 662 TTACCATTAAGATTAAGATTTATCATCCGGGAGCTGGAAGATCTGGAAATGATCTT 721
|
|
|
Db 3158 TTACCATTAAGATTAAGATTTATCATCCGGGAGCTGGAAGATCTGGAAATGATCTT 3217
|
|
|
QY 722 CAGTAGTAAACAAACAGCAAGAAAGCAGAGACCATTAAGGGAAGAAATAGATGACTAC 781
|
|
|
Db 3218 CAGTAGTAAACAAACAGCAAGAAAGCAGAGACCATTAAGGGAAGAAATAGATGACTAC 3277
|
|
|
QY 782 AGAGAGAGCTGATATATATTTAATTAACCTTAGTTCTGAACTCATTCGCGCATGTGGGAGC 841
|
|
|
Db 3278 AGAGAGAGCTGATATATATTTAATTAACCTTAGTTCTGAACTCATTCGCGCATGTGGGAGC 3337
|
|
|
QY 842 CTGATTAACCCATTTGTCAGAAAGATATACATGAGTTAAATTCAGGATGGGATTTCTTA 901
|
|
|
Db 3338 CTGATTAACCCATTTGTCAGAAAGATATAGAGTAAAGT----- 3370
|
|
|
QY 902 AATTAAGCTTGAAGAACCCGATTTGACCAACTTGAAGCAATGACGTCGCTCAGTACA 961
|
|
|
Db 3371 ----- 3370
|
|
|
QY 962 GATGACCTGACGCGGTATTTTCTGGGTAGATATTTGACGCTGTGATGCTTCCGTTAATG 1021
|
|
|
Db 3371 -----AGCGGTATTTGACCTGGGTAGATATTTGACGCTGTGATGCTTCCGTTAATG 3420
|
|
|
QY 1022 TCTCCAAATGGAACAGCATCTGGAACCTGCAAGCAGCATTTGATGCTGAGCAAT 1081
|
|
|
Db 3421 TCTCCAAATGGAACAGCATCTGGAACCTGCAAGCAGCATTTGATGCTGAGCAAT 3479
|
|
|
QY 1082 TAACTGAGGCTTATCAACAGCAGATAG-AGTGTACAGACTGACTCATCAGCAGACGC 1140
|
|
|
Db 3480 TAACTGAGGCTTATCAACAGCAGATAG-AGTGTACAGACTGACTCATCAGCAGACGC 3538
|
|
|
QY 1141 TTTTGTAAAGAAAGTAAACAGAGAGTGAACAACACTGTTCTAGACCCAAATAATG 1200
|
|
|
Db 3539 TTTTGTAAAGAAAGTAAACAGAGAGTGAACAACACTGTTCTAGACCCAAATAATG 3598
|
|
|
QY 1201 AACTGATTTGATAT-GGCTAGCCTGGAAGAGATCATCAACAGACGATTAACCTGG 1259
|
|
|
Db 3599 AACTGATTTGATATGGGATAGCTGAGAGAGAGATCATCAACAGACGATTAACCTGG 3658
|
|
|
QY 1260 AAGGTGCTCTATTAGCCNTGGGTCAAGTTCCAAATGCTCTGATAGCTCTG--CANTG 1317
|
|
|
Db 3659 AAGGTGCTCTATTAGCTTGGGTCAAGTTCCAAATGCTCTGATAGCTCTGAGCATGGC 3718
|
|
|
QY 1318 TGAACACACCGAGGGCTTGTAAAGTGAAGCAAACTGTTGGAGAGACCTTAAGCCA 1377
|
|
|
Db 3719 TGAACACACCGAGGGCTTGTAAAGTGAAGCAAACTGTTGGAGAGACCTTAAGCCA 3778
|
|
|
QY 1378 TTGAATTTGAATGCTGCAAGCATCATGTGCTCAAAATGATTAAGCCATCGTCA 1437
|
|
|
Db 3779 TTGAATTTGAATGCTGCAAGCATCATGTGCTCAAAATGATTAAGCCATCGTCA 3838
|
|
|
QY 1438 CAGTGAAGCCGTTAATAAGCAGAGAAATGATCTAATTTGAATCAAGTGCAGAGAGAAAG 1497
|
|
|
Db 3839 CAGTGAAGCCGTTAATAAGCAGAGAAATGATCTAATTTGAATCAAGTGCAGAGAGAAAG 3898
|
|
|
QY 1498 CAAGCAACCTTCAGAACCAAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTGGAAA 1557
|
|
|
Db 3899 CAAGCAACCTTCAGAACCAAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTGGAAA 3958
|
|
|

ARRAKQFHEBANSKIMBWESEKSLDSELEIANDPDKIKAOLVQHKEROKSLGKHS
VDTNRGSLKLEKTSLADNDLKNMSELBDKWDITCGKVEONKMLEBLTSG
OFEDALQALIDMLRYVEPOLAEDQPHYDIDLVMSLDHVKQKELGKRSVLTG
RSARELESGRSDSWVRVQOMELSTRMEVICALSLSKORLESLKQAEBSVYHT
LMELEABOTLBFHGLPDDEDALRTLLEONKEPKRLEEKRAELSKATQEDALLA
VCHPDSITTTIKMTITTIQARFEVILAMAKHOORLAGLAGLAKOBLLETLIAWLO
AETTLKEDKEVIPOSEBVKTLIAHQTFMEBMTKQDPVDVTKYTKRATDPSSL
OSHIPVLDKGRARKEPASFYPSGSQTOIETKNRVMVLVSKWOQWMLALERBK
LNDALRLBELREFANFDPDIWRKTYRMWNHKSVMFPRF1IDDOGKITROEPI
DGLISKEPFLREMSAVADIIPRDGQYDIYEFVALHEDKDAYKAPITDADKIGDE
VTRQVAKCKAKRFOYBOIGDNKRFELFNQPGDSOOLRVITLSEITWVRVGGKMA
LBEFLYKNDPCRVHNGSKMLREBSNSSTIATOPTLAKRTIMELBKFLIADGASOG
MAAFPRGRSRSPRSNGASPNKSTSSSHACQASPPVPAASVTPTKGTQSKLPL
GVYSGFHSGBESDALITTAARVKTQFESRKTSPRPGSRASKASRSRSGDA
SPDISEIOSVCSDETVPTHRPVRPAGSRPSTAKPSKIPTQRSSPASKLDSKR

ORIGIN

Query Match 58.6%; Score 1190.8; DB 10; Length 17206;
Best Local Similarity 80.3%; Pred. No. 2.6e-278;
Matches 1507; Conservative 0; Mismatches 351; Indels 18; Gaps 9;

4 CCATGCTGACAGAAACAAATCAATCATCTCAGCT-CCGGCCCCAGCCCTTGAAATAG 62
DB CTTGGCTTAACGAAACAAACGAATCATCTCAGCTGCTGCCCCAGCCCTTGAGATAG 11350
63 ATACTTAAGGACAGACAGAGAAAGAAATCGGCAATGCGTGAATTGATAGTGAACACA 122
DB AGACTTTTGAAGACAGCAAGAGAGACCGGCAACTTGGAGAACTGATAGCCGAGACA 11410
123 AGCTCATATGATTAAGTGAACAAAATGCG---CCACATGACTGAATGAGCCCTGGGG 178
DB AGCACACATAGATGAATGATGACAGCTGGGACAGCTATGGAATAGCCCAAGG 11470
179 AAGCTTTTCTATCCAGAGAGATGTTGGCAGCCGACACCTTTACGTCATGATTAAG 238
DB AAGGCTTTATATCAAGAAAGATAGTGGACGCGACCTCTTACAGTCAATTAAG 11530
239 AAGATGTCAAAAACGCTGTGTGGCACTGGATGAAGCAATTTCTCAATCAACTCAGTTC 298
DB AAGATGTCAAAAACGCGCTGTGTCTTGAATGAGCAATTTCTCACTCAACCAAGTTTC 11590
299 ATGACAAGATGATCGAGTCTTGAAGAGCTTGAAGCGCATGTTGAACGTTGAAGCAGC 358
DB ATGACAAGATGACAGATCTTGAAGAGCTTGAAGCGCATGTTGAAGCGCTGCGCAGC 11650
359 CACCTCTATCTGTGAGAGGTTGAAGAGATCAAGAGACGATCAAGTGAAGAAATTAAGAG 418
DB CGGCATCTCATCTCGCGAGGTTGAAGAGATCAAGAGACGATTTGTAAGAAATTAAGCGC 11710
419 TGTCAGTAGACATGAGAAAGCTACAGCCGTTGTATGAACTCTTAAACAGAGGGAGAG 478
DB TGTCCGTGACATGAGAAAGCTACAGCCGTTATACAGACTCTGAGGACAGAGGGAGAG 11770
479 AATGATTTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAAGCTGTTCAAGATA 538
DB AGATGATCGCAGATCCGAGGGACATGAAAGACGATCTGCAGAGCTGTACAGATA 11830
539 AGCTTACCAAAAGGTTTTCAATTTGGAGAAACATPACACACATGTTGGAAGAGAGAGAG 598
DB AGTTGACCAAAAGGTTTTCACTTGGAGAGACATCCACACGCTGTGGAAGAGAGAGAG 11890
599 CCAAACTAGTGAATGATGAGACTAGCAGAAAAAGTTGTGGTGTGATCAATGTCAATGGA 658
DB CCAAACTAGTGAATGATGAGACTAGCAGAAAAATTTCTGTGTGATCAATGTCCGTGG 11950
659 TAGTTACCATTAAGATCTCAAGATTTTCACTCCGGGACCTGGAAGATCTGGAATTGATC 718
DB TGTCTACCATTAAGACACTCAAGATTTTCACTCCGGGACCTGGAAGATCTCCGAATGACC 12010
719 CTTCAATGATTAAGACAG 778
DB CTTGGGTGTAAG 12070

779 TACAGAGAGAGCTGATATATAGTTATTAACCTAGTTCGAACTCATTTGGCGGATGTGGG 838
DB TGCAGAGAGAGAGCTGACATGTGATCATACCGCTAGGCTCAGAGCTCATCGGTGATGTGGG 12130
839 AGCTGATTAACCACTTGTCAAGAGAGATATACATGATTAATTAAGCGATGGATTTCT 998
DB AGCTTACCAAAACCACTTGTCAAGAGAGATTAACAGATTAATTAATTCAGC-ATGGGATTTCT 12189
899 CTAATAAAGGCTTGAAGACCGGATTTGACCACTTGA-----GAGCAATGACGTGCC 952
DB CTAATAAAGCTTGAAGAGACCGGATTTGACCACTTGAAGAGAGAGAGAGAGAGAGAGAG 12190
953 GTCAATGACAG 1012
DB CAGTACAG 12309
1013 GCTTCAATGTCTTCAATTTGGAACAGATCTCGAAACTGTTCAGACAGAGAGAGAGAGAG 1072
DB GCTTCAATGTCTTCAATTTGGAACAGATCTCGAAACTGTTCAGAGAGAGAGAGAGAGAG 12310
1073 AGAGCAATTTAAGTCTGAGGCGCTATCAAGAGAGATAG-AGTGTGACAGCTGATCA 1131
DB GAGCAATTTAAGTCTGAG 12428
1132 CGCAGAGCTTTTGTCTTAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
DB AGCAGA-ACGTGTTGCTGAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12487
1192 CAAAAATGAGACTGATTTGATAT-CCGATGCTGGAAGAGAGATCATCAAGAGAGAG 1250
DB CGCTGATGAG 12547
1251 ATAACTGAG 1310
DB ACAAGTCTGAG 12548
1311 T-GCATGCTGACACACACAG 1368
DB TGCATAGGCTGACACACACAG 12667
1369 CTAAAGCCATTGAATTTGAATTTGCAAGATCATATGCTCCAAAATGATATTAAGGCC 1428
DB CGAAAGCCATTGATTTGAATTTGCAAGATCATATGCTCCAAAATGATATTAAGGCC 12727
1429 ATAGTCCACAGTGAAG 1488
DB ACCAGTCCACAGTGAAG 12728
1489 GAG 1548
DB GAG 12788
1549 TTTTGAAGAAACAG 1608
DB TTTTGAAGAAACAGAGATCAAG 12848
1609 GGTTCATGAG 1668
DB GGTTCATGAG 12908
1669 TGGCATCTAAACGCTGGAGAGTTTACCGGAAACAGCCAGAGAGAGAGAGAGAGAGAGAG 1728
DB TGGCATCTAAACCTCTGGAGAGTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12968
1729 TGGAGTCTGTGCTGCTTTGAAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1788
DB TGGAGTCTGACCGGCTTTGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13028
1789 GCCAGAGATGCTTGAAG 1848
DB GGCAGAGATGCTTGAAG 13088
13088 GGCAGAGATGCTTGAAG 13147

Qy 1849 ATAACTGAAAAAAA 1864
Db 13148 CCAACTGAAAGAGA 13163

RESULT 9
AF396879
LOCUS 22218 bp mRNA linear ROD 21-SEP-2001
DEFINITION Mus musculus bullous pemphigoid antigen 1-b (Bpag1) mRNA, complete
ACCESSION AF396879
VERSION AF396879
KEYWORDS AF396879.1 GI:15077864
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Leung C.L., Zheng M., Prater S.M. and Liem R.K.
TITLE The BPAg1 locus: Alternative splicing produces multiple isoforms with distinct cytoskeletal linker domains, including predominant isoforms in neurons and muscles
JOURNAL J. Cell Biol. 154 (4), 691-697 (2001)
MEDLINE 21405767
PUBMED 11514586
REFERENCE 2 (bases 1 to 22218)
AUTHORS Leung C.L., Zheng M., Prater S.M. and Liem R.K.H.
TITLE Direct Submersion
JOURNAL Submitted (05-JUL-2001) Pathology, Columbia University, 630 W. 168th St., P&S 15-421, New York, NY 10032, USA
FEATURES
SOURCE location/Qualifiers
1..22218
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="1"
/cissue_type="muscle"
1..22218
/gene="Bpag1"
13..22182
/note="cytoskeletal protein BPAg1"
/codon_start=1
/product="bullous pemphigoid antigen 1-b"
/protein_id="AAK8384.1"
/db_xref="GI:15077865"
/translation="MAGYLSPAAYVVEBOEYLAAYEDVLERKYDERDKYOKKFTKM
INQIMKRYKRVNDLYEDLDNGHNLISLEVLSDPTPREKGRFRFRLLQVQALALY
LKRQVKLVIRNDITDGNPKLTGLIWTILHFQISDIHVTESEDSMAKERLLIM
TQOATEGAVACRENTTCWRDGLFNALIHKYPDLDIMNTAVQSNLNLERAFV
AEKIGVRLDPEEDVDSPEDEKVLIVSLDAPKPVPGEGIGANDVEMKIFY
OMNMYLIOMIRHVVMTSERTEFPNNLELALNOYLOFKEKIEPREMEKXIKEL
YKLEIWEIEGRITLGGYHNDIEKEMGKLTIMLEKALPEVERLMLCOIAR
VQDSVSCEDKILARNALQSDSKRLSSQVQFQWBAIAGIIECENLROHTDVI
LIDKRYQAOQLVQVAKLRDEIHALNECSVYSKRMLTTEGTQKIMISGITQSLNS
GPAQTLPINSNGTOSLTPTSLTSSVTSGLSSGMTSLTPSPVPAVPGSVVAAN
FSLGVEENSLQTLKMOIRKPLKSLDQNLTEEVNMMKFDQLMMVDMQVQDLR
TEMGSDLPVSESLHNHNVRAIEEPSSLKEAKISEIOMTAKLKYNDKHLRLS
OYKLTNTSRNOERHLDLTHNFUTRALNTLWNEKESSEAYVMSRATYMSRQCKY
HVELMLEEQEESIKAYOEIAEOLILNHPARLTITAYRANMOTQNSWILQCCYE
OHIQENSAYPEFENDAEADYLNLDALQKYSCDRSSIHLLEDLVQESMEKEEL
LOYRSVAGLMSRAKTVVOLKPNNDPKLTSIPDKAIQDROIEITTYDDCEVLN
NSHRAKKVLSPTENEAVPSVCTPPPKAEAVDFANRLEQQOQSVLTLMHESHIM
KSYSMTLVNEIDIRASNVASIKTMLPGHQOVLNLSGRLEDPLDESGOEIFSG
SDISOLEKESVCRKYQOELIKSAREBOESVNYLTVSEPNRLRLSEGEDLITRO
IKRPLESDDLHESMLRTTBQKLEKELDKODIGTTNKCEPFSQAABSPPYAIR
SELSVVIQSLQISMSSTYIEKLTANVLVKNQAABALVKYTEKLCCEBAIYAK
NNTENLSTLKQMSRSEVDEKREPHALEDLQAKAISDEMFHKERDIDFWMHK
ADOLVERMOSVHVQIDNRLDLEIGKSLKHVDSYHPLDQWIEITTOIKIQENP
ENSKALALQNLQKLVSEIKQSKDECKYQSEYSAVKDYLQMTYRAVQESQ
OKSPVRRRIQSSADLVIOEFMDRTTYTLVLTMTQYIKRPAQGISLREBEKSLDE
EKQHIIEKAELOKRVNSIKITLDGERAKGRPLFSKOQMSKEISTKKEQSEALQTT

[illegible]

ORIGIN

Query Match	58.6%	Score 1189.2;	DB 10;	Length 22218;
Best Local Similarity	80.3%;	Pred. No. 6.6e-278;		
Matches 1506; Conservative	0;	Mismatches 352;	Indels 18;	Gaps 9;

OY	4	CCATGCTGACGAAACCAATCAATCATCTCTCAAGT - CCGGCCAGCCCTTGAATATG	62
Db	17318	CTTGCGCTACGGAAACCAACGATCATCTCTCAAGTGGCTTGGCCAGCCTTGAGATGCG	17377
OY	63	ATACTTAAGGCGACGACGAAAGAACATGGGCACTGGCTGATGTGATAGCTGAACACA	122
Db	17378	AGACTTTTGGAGACACACAAAGAGAGACCGGCAACTCTGAGAACTGAATAGCGAGACA	17437
OY	123	AGCTCATATAGATTAATGAACAAACCTGG --- CCACAGTACTGAATAGAGCCCGGGG	178
Db	17438	AGCCACACATAGATTAATGAACAGCTGGGCCACAGCTATTGGAACTGAAGCCAAAG	17497
OY	179	AAGCCTTTTCTATCCAAAGAAAGTATGTGGCACCGACA C C C T T T A C A G T C A C A T T T A A G	238
Db	17498	AAGCATTATATCCAAAGAAAGTATGTGGCACCGACA C C C T T T A C A G T C A C A T T T A A G	17557
OY	239	AAGATGTCAAAAACCGTGTGTGGCACGTGATGAGCCATTTCTCAATCAATCACTCAAGTTCC	298
Db	17558	AAGATGTCAAAAACCGAGTGTGTGGTTCTGATGAGGCAATTTCTCACTCAACCAAGTTTC	17617
OY	299	ATGACAAAGTATGATCAGATCCTTTGAGAGCCTGGAAACGATCGTGGAAAGCTTGAAGCGAC	358
Db	17618	ATGACAAAGATCGACCAATCTTTGAGAGCCTGGAAACGATCGTGGAAAGCTTGAAGCGAC	17677
OY	359	CACCCCTATCTCTGCGAGAGGTTGAGAAAGTTCAAAGAACAGATCAATGAAATTAAGAATG	418
Db	17678	CGCCATCATCTCCGCGAGAGTGAAGAAAGTTCAAAGAGAGATGTGTGAAATTAAGAGCG	17737
OY	419	TGTCAGTAGACATGGAANAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGGGAGAGG	478
Db	17738	TGTCCTGGACATGGAABAAGCTGACCCGTTATACAGACTCTGAGGCAAGAGGGAGAGG	17797
OY	479	AAATGATTTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAGCTGTTTCAGATA	538
Db	17798	AGATGATCGCGAGATCCGAGGGCACTGAAAAAAGCATTCTGCAAGCTGTATCAGATA	17857
OY	539	AGCTTGACCAATGTTTTCAATTTGGAGAACATACACACTGTGTGAAGAGAGGAG	598

Db	17858	AGTTGGACCAAATGGTTTTCATCTGGGGGAAGCATCCACACGCTGGTCTGAAGAACGGGAAG	17911
Oy	599	CCAAACTACTGATGTGATGATGAGACTAGCAAAAAGTTCGTGTGATCACATGTCAATGA	658
Db	17918	CCAAACTGCTGGATGTGATGATGAGACTAGCGGAATAATTCGTGTGCATTCACATGTGCTGG	17977
Oy	659	TAGTTACATTTAAAGATACTCAAGATTTTCATCCGGGACTCGGAAGATCCTGGAAATTGATC	718
Db	17978	TGCGCACCATCAAAAGACACTCAAGATTTTCATCCGGGACTCGGAAGATTCGGGAATTCGACC	18037
Oy	719	CTTCAGTAGTAAACAAACAGCAAGAAAGCAGAGACCAATTAAGGGGAATAATGATGAC	778
Db	18038	CCTGCGTGTGTAAGACAGCAAGAAAGCTGCAAGGCAATCAAGAGGAGATTTAGCGCC	18097
Oy	779	TACAGAGAGACTGATATATAGTATTAACTAGTTCAGTCACTGCGGATGTGGG	838
Db	18098	TGCAGAGAGAGCTGGACATGTGATCAAGCTTAGCTCAAGGCTCATGCTGATGTGGG	18155
Oy	839	AGCTGATAAACCCATTTGTCAGAAGAGTATACATGATTTAAATTAGCGATGGATTCT	898
Db	18158	AGCTGCAAAACCCATTTGTCAGAAAGAGATTTGACAGCTAAATTCAGC-ATGGGATTCCT	18216
Oy	899	CTAAATTAACCTTGGAAAGACCGGATTTGACCACTTGA-----GAGCAATGCAGCTGCC	952
Db	18217	CTAAATTAACCTTGGAAAGACCGGATTTGACCACTTGA-----GAGCAATGCAGCTGCC	18276
Oy	953	GTCAGTACATAGGACTGACAGCGGATTTTCTGGGTAGATTTCAGAGTGTGAATTC	1012
Db	18277	CAGTACCAAGACGGATTTGCAGGGAAATTTTCGACTGGTTGACATCGCAGGCGACAAGTTA	18336
Oy	1013	GCTTCAATGTCTCCAAATTGGAAACAGCATCTGAACTGTCAAGCAGCAGATTTGATGTCT	1072
Db	18337	GCCACATGTCTCCCAATTGGAACAG-ACCTGAGAGCCGTCAAGCAGCAGATTTGAAGACT	18395
Oy	1073	AGAGCAATTTAACTCTGAGGCTCATTAACAGCAGATAG-AGTGTGATCGATCACTCATCA	1131
Db	18396	GAAGCAATTTAACTCTGAGGCTCATTAACAGCAGATAGTGAAGTGAAGGCTGAACCATCA	18455
Oy	1132	CGCAGACGCTTTTGCTAAAGAAAGTAAACAGNAGAGTACAAACACATGTTCTTAGACC	1191
Db	18456	AGCAGA-ACGTGTGCTGAAGAAAGTAAACAGAGAGGCAACACAGCACACAGTCCAGACC	18514
Oy	1192	CAAAATGGANCTGATATTGTAT-T-GGCTAGCCTGGAGNAGAAATCATCAACAGACGC	1250
Db	18515	CGCTGATGAGCTGGAAGTTGATCTGGGACAGCCTGGAACAGGAATTTGCACAGACAGC	18574
Oy	1251	ATTAATCTGAGGGTGCCTTAATTAGCCNTGGGTCACTTCCAATGCGCTGTATAGCTCC	1310
Db	18575	ACAAGCTCGAGGGTACTCTCTTAGCACTGGGTCAAGTTCCAGACAGCGCCCTGACAGACTCC	18634
Oy	1311	T--GCATGCGACACACCGAGGGGCTGTAACTAGAGAGAAACCTGTGGAGAGACC	1368
Db	18635	TGCGATGCTGACACACCCAAAGGCTTGTAACTAGAAACAGAAACCCGTTGAGGAGACC	18694
Oy	1369	CTAAAGCATTGAAATTTGAATTGCTCCAGACATCATGTGCTCCAAATATGATTTAGGCC	1428
Db	18695	CGAAAGCATTTGAGATTGAATCTGGCTAAGATACGTGCTCCAAAACAGACGTCTTAGCCC	18754
Oy	1429	ATCAGTCCACAGTGAAGCCGCTTAATTAAGCAGGAATATGATTAATTGATCAAGTCAAG	1488
Db	18755	ACCAAGTCCACGATGAAGAGCCGTCAACAAAGCAGGAATATGATCTCATCGAGTCTGAATGAAG	18814
Oy	1489	GAGAAAGAGAACCACTTCAGAAACAACCTAGAGGTTTAAATCAACGCTGGCAAAATNG	1548
Db	18815	GCAGAGAGCAACCACTTCAGATCAACCTCAGAAATTCATATCAATCAGCCTGGCAGGATTA	18874
Oy	1549	TTTTTGGAAAAACAGAAACAAAGGAACACAGCTGATGTCTCTTGCCACAGGCCAAAG	1608
Db	18875	TTTTTGGAAAAACAGATCAAAAGAAACACAGCTGACAGCCCTTGCCACAGGCCAAAG	18934
Oy	1609	GGTTCCATGGCAGAAATGAGATTTTGCAGCAGTGTGCTGACACAGAGCGTCATCTGT	1668
Db	18935	GGTTCCATGGTGAATTCAGAAATTTTSCACAGTGTGTGCGGACACAGAGCGTCATCTCT	18994

Qy	1669	TGGCATCTTAAACCCCTGGAGGTTTACCGGAAA	CAGCGAAGGAGCAGCTTAAATGTCATA	1728
Db	18995	TGGCATCTTAAACCTCTGGAGGCTTCCACAGAC	ACGCCAAGGAGCAGCTTAAAGCACATA	1905
Qy	1729	TGGAACTCTGTCGCTGCCTTTGAAAGCTAAAG	AAGAAACATATAAGATCTGATCAGAAAG	1788
Db	19055	TGGAACTCTGCACCGCTTTTGCCATCAAGGAG	AGACGTCACAAAGCCTGATCTTAGGG	19114
Qy	1789	GCCAGCAGATGCTTTGCAAGATGCCCAAAAT	CTGCAGAGCAAAATATTGACCAAGACATA	1848
Db	19115	GGCAGCAGATGCTTTGCAAGGATGTCGCCGAT	CTGCGAGGACCAACATTGACCAAGACATCA	1917
Qy	1849	ATACTTGAAGAAAAA	1864	
Db	19175	CCAACTGAAAGAGAA	19190	

RESULT 10					
AF083131	AF083131	1022 bp	mRNA	linear	PRI 10-AUG-2000
LOCUS					
DEFINITION	Homo sapiens CAVX-15 mRNA, complete cds.				
ACCESSION	AF083131				
VERSION	AF083131.1	GI:9789026			

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS Kairo, A., Wang, L., Gao, Z.Q., Gao, Z.P. and Boman, B.M.
TITLE Isolation of Novel Genes From Human Colonic Epithelial Cells

REFERENCE	AUTHORS
2 (bases 1 to 1022)	Kalro, A., Wang, L., Gao, Z. Q., Gao, Z. P. and Boman, B. M.

TITLE Direct Submission
JOURNAL Submitted (10-AUG-1998) Cancer Research, CA*TX, Inc., P.O. Box 710,

FEATURES	Location/Qualifiers
source	1. .1022

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/tissue_type="colon"
```

gene	1.1022
	/gene="CATX-15"
CDS	166.906

```
/gene="CATX-15"  
/codon_start=1  
/product="unknown"
```

```

/protein_id="AAFA98241.1"
/db_xref="GI:9789027"
/translation="MSPGEGGFSIC

```

HDKIDILLESIRIVEELRQPPISIAEVEKIEKQISENNKVSVMDEKLOPIYETLKQR
GEMTARSGGTDKISAKAVQDKLDQWFIWENTHTLVEBEARKLDMELAEKFCND
HMSLVTIKCTODEFRDLEPGISVYKOOOEAETIRREIDGLOEELDIIVNIGSSE

ORIGIN

Query Match	50.3%	Score 1022	DB 9	length 1022
Best Local Similarity	100.0%	Pred. No. 2.2e-237		
Matches 1022; Conservative	0	Mismatches	0	Gaps 0

1 TTGCGATGCTGACAGAAACACATCATCATCTCTAGCTCCCGCCCCAGCCCCCTTGATA 60

61 TGATACTCTTAAGGCACGACAGAGAAGACATCGGCMACTGCGTGAGTTGATAGCTGAACA 12

Db 61 TGATACTCTAAGGCAGCAGCAGGAACAATCGGCACTGCGTGAGTTGATAGCTGAACA 12

Db 121 CAAGCTCATATAGATGAACAAACTGSCCAGTACTGATGAGCCCTGGGAA 180

[illegible]

241 GATGCAAAAGCGTGTGGCACTGGATGAAGCCATTTCATCAATCACTCAGTTCCAT 300

Db 241 GATGCAAAAAGCGTGTGTGGCAGTGAAGCCATTCTCAATCAACTCAGTTCCAT 300

Db 301 GACAAGTAGATCAGATCCTTGAGACCTGGAACGATCGTGGAAAGCTTGAGGCAAGCA 360

361 CCCCTATCTCTCCAGAGGTTGAGACATCAAGCAACAGATCAGTGAAATAAGATGTG 420

421 TCAGTAGACATGGAAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGGGAGAGGAA 480

Db 421 TCAGTAGACATGAAAAGCTACAGCGTTGTATGAACTTTAAAACAGAGGGGAGAGGAA 480

481 ATGATTGCTAGATCTGGGGGAACTGATTAAGAGATATCTGCGCAAGCGCTTCAGAGTAAG 540

541 CTTCACCAATGCTTTTCA TTTGGAGAACATACACACTGCTGGAAGAGAGGGAAGCC 600

541 CTTCGACAAATGGTTCATTTGGGAGAACATACACACACTGGTGGAGAGAGGGAAGCC 600

Db 601 AAACACTGAGTGTGATGAGCTAGACAGAAAAGTTCTGTTGTGATCACATGTCATTGATA 660

QY 661 GTTACATTAAAGATACCAAGATTTCATCCGGGACCTGGAGATCCCGGAATTGATCCT 720

721 TCAGTTAGTAAACACACAGCAGAGCAGACCATTAAGGAGAAATAGATGCACTA 780

Db 721 TCAGTAGTAAACACAGCAAGACGACAGAGCCATTAAGGGAGAAATAGATGCACTA 780

781 CAGGAGGAGCGGATATAGTTTAACTAGGTTCTGAACCATTCGGGCATGCGGGAG 840

841 CCTGATTAACCCATTGTCAGAAGAGTATACATGAGTTAATTACGCCATGGGATTCTCT 900

```
001      XXXXXXXXXXXXXXXXXXCCCCCCCCCCXCCXXCXCXCCCXCCCCCXCCXCC  
DB    CCIGAIAPACLCCHGICLCAHGAAGAGTATCACATGCAGI   IAAAXI   CAGCGAII GGGAILCLCI     900
```

Db 901 AATTAAGCTTGGAAAGACCGGATTGACCACTTGAGAGCAATGCAGCTGCCCTCAGTAC 960

961 AGATGACTGCAGCGCGTATTTTTCGGGTAGATAATTCAGSTGGTAACTCCGTCAT 102

QY 1021 GT 1022

Db 1021 GT 1022

RESULT 11
CQ729862

DEFINITION	Sequence 15796 from Patent WO02068579.
ACCESSION	C0729862

VERSION	DATE
KEYWORDS	.
SOURCE	Homo sapiens (human)

Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS
TITLE Venter, C.J., Adams, M.C., Li, P.W. and Myers, B.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcript, for detecting expression and other uses
therof
Patent: WO 02068579-A 15796 06-SEP-2002;
JOURNAL
PE Corporation (NY) (US)
Location/Qualifiers
1. 15293
FEATURES
Source /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 33.2%; Score 674.4; DB 6; Length 15293;
Best Similarity 63.2%; Pred. No. 9.3e-153;
Matches 1134; Conservative 0; Mismatches 645; Indels 15; Gaps 6;
QY 41 CCGGCCCGCCCTTGAAATGATCTAAGCAGCAGCAGAGAAACATCGCAACTG 100
DB 10554 CCTCTCCAGCCATTGATGATGAGCAGCTCAGCAGCAGCAGAAATGAGGCAATTA 10613
QY 101 CGTGAAGTTGATGCTGAACCAAGCTCATATGATGATGAACTGGCC---A 156
DB 10614 AGGGAATCTATGCTGCAACCAACCTATATGACAACTACTAAAGATAGGCCCA 10673
QY 157 CAGTACTGATGAGCCCTGGGGAAGCTTTCTATCCAAAGAAAGTATGAGCAGCCGAC 216
DB 10674 CTAAAGAAATTAACCTCGAGAAAGGGAATGATGAAATAATTCAGAAAGCAGAA 10733
QY 217 ACCCTTACAGTACATTAAGAAGATGTCAAAAGCGTCTGTGCACTGATGAAGCC 276
DB 10734 AACATGATGCCAAATTAAGAGAGAGTGGCCAGAGAGCCCTGCTGATGAAGCC 10793
QY 277 ATTTCTCATCACTCAGTTCCATGACAAATGATGATCCTTGAAGCCTGGAAGCC 336
DB 10794 GTGTCCAGTCCACACAGTTTCAATGATTAATGAGCTATGTTGAGACACTGGAAT 10853
QY 337 ATGCGGAAGCTGTGAGGAGCAGCCCTATCTCTGCAAGGTTGAGATCAAGAA 396
DB 10854 CTTTCTCTGCTGCTGCTGATCCACCACTGATCCCTGCTGAAGTGAAGTCAAGAG 10913
QY 397 CAGATCAGTGAATAATGAATGTGTGATGATGATGAGAAAGCTACAGCCGTGTATGA 456
DB 10914 TGCATCAGTGAATGAAGATGCGACCGTGAAGTGAATAATTCAGCATCTTTGAG 10973
QY 457 ACTCTTAACAGAGGGGAGAGAAATGATGATGATGATGATGATGATGATGATGAT 516
DB 10974 GCTTGAAGCGCGTGAAGAGAGCTTATGAGAGATCTCAGGAGAGCAGCAAGATCTG 11033
QY 517 TCGCCAAAGCTGTTCAGATTAAGCTTGACCAATGTTTTCATTGGGGAACATACAC 576
DB 11034 GCTGCAAAAGAAATTCAGATTAATGATCAATGATGATGATGATGATGATGATGAT 11093
QY 577 ACACTGATGAGAGAGGAGGAGCAAACTACTGATGATGATGATGATGATGATGATGAT 636
DB 11094 GCTCGGGCTGAAGAGAGAAATCAATTTCTTGATGATGATGATGATGATGATGATGAT 11153
QY 637 TGGTGTGATCAGATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 696
DB 11154 TGGTATGACATGTCAGCTCTGACACCATCAAAAGACACCCAGGATTTGTCATGAC 11213
QY 697 CTGGAAGATCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
DB 11214 TTGGAAGAGCCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11273
QY 757 ATAAAGGAAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
DB 11274 ATTAAGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11333
QY 817 GAACTCAATGCGCATGTTGGGAGGCTGATTAACCCATTGTCAAGAAAGATATACATGAG 876

DB 11334 GATTGATTTTTGCTGTGAGAGAACTGAGAACTGAAAGTGAAGAGATGATGATGAG 11393
QY 877 TTAATTCAGCATGAGGATTTCTTAATAAGCTTGAAGAGACCGGATGACCAACTGA 936
DB 11394 ATGAATTAATGC-TTGGAGAGACTTAACAAACATGAGAGAGAGCTAGAAAACTTGA 11452
QY 937 G-----AGCAATGACAGTCCGCTCAGTACAGATGATGATGATGATGATGATGATGAT 990
DB 11453 GATGCTATGCAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 11512
QY 991 AGATATGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
DB 11513 AGATTAACATGTGATTAATTAATCTGACACATGCCCCCTGTGGACATG-ACCTCAATATCTG 11571
QY 1051 TCAAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110
DB 11572 TTAAAGATCAGTTAAATGAATGAAGAGATTTCAAAGTGAAGTTTACCAACCAAAATTTG 11631
QY 1111 AGTGTATGCACTGATCATCAAGCAGCAGCAGCTTTTGTAAAGAAATGAACAGNAGAGT 1170
DB 11632 AGATGAGAGAGCTTAATCAACAGGATGAATGATGATTAAGAAAGCTATGATGAGACGG 11691
QY 1171 ACAACACACTGTTTGAACCCAAAAATGAGANTGATATTTGATAT-GGCTAGCTTGAGN 1229
DB 11692 ACAGAGACATTAATACAGAGAACCTGACAGAACTCAAAACCTCTGGAGAACCTGGGTG 11751
QY 1230 AGAGATCATCAACAGACAGCAATAAAGTGAAGGCTCTGATTAAGCCTGATGATGATGAT 1289
DB 11752 AGAAATGCTCCACACAGCAGCAGCAAACTGAAGGGCTCTGTGGCCCTTGTGATGATGAT 11811
QY 1290 AACATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
DB 11812 AGCATGCTTTAAGAGAACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 11871
QY 1348 AGAAACCTGTTGAGAGAGACCTTAAGCAATGATGATGATGATGATGATGATGATGATGAT 1407
DB 11872 AGAGACCAATTAAGTGAAGAGACCAAAAGTCAATGATGATGATGATGATGATGATGAT 11931
QY 1408 TCCAAATGATGATTAAGCCATCACTGACAGTGAAGCCGTTAATGAAGCAGAAATG 1467
DB 11932 TAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11991
QY 1468 ATCTAATGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1527
DB 11992 AGCTTGTGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12051
QY 1528 TAAATCAAGCTGTGCAAAATGTTTGAAGAAAAAGAAAGAGAGAGAGAGAGAGAGAG 1587
DB 12052 TGAACCAATGCTGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 12111
QY 1588 GTGCTTGGGCGCAGGCAAGGGTTCATGGGGAATGAGATTTGACAGCAGTGGCTGA 1647
DB 12112 CAATCTGACAGGCGCCAGGCTTCCACATGATGATGATGATGATGATGATGATGATGAT 12171
QY 1648 CTGACAGGAGCGCTCATCTGTTGGCATTAACCCGCTGGAGGTTTACCGGAAACAGCA 1707
DB 12172 CTAGATGAGAGAGCAGCTTTCTGATCTTAAGCCACAGAGAGACTTCCGAAACGCTGA 12231
QY 1708 AGAGAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1767
DB 12232 GGGAAAGCTTGAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 12291
QY 1768 ATTAAGAGTCAATGAGAAAGGCGAGCAGATGTTGCAAGAGTCCCAAAATCTG 1821
DB 12292 ATTAATCACTTGAAGAGGAGAGCTCATGCTTGAAGCGGTGACGACTCTG 12345

RESULT 12
AX677291 6007 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 69 from Patent WO02086122.
ACCESSION AX677291
VERSION AX677291.1 GI:29334696

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1
Legrain, P. and Davlet, L.
Protein-protein interactions in adipocytes
Patent: WO 02086122-A 69 31-OCT-2002;
JOURNAL Hybrigenics (FR)

FEATURES
Source Location/Qualifiers
1..6007
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 32.3%; Score 655.4; DB 6; Length 6007;
Best Local Similarity 62.9%; Pred. No. 3.5e-148;
Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;

QY 41 CCCCCCCCCCTTGAATATGATCTTAAGCGCAGCAAGAAACATCGCACTG 100
DB 1259 CCCCCTCCAGCATTTGATCATGAGCAGCTCAGCGCAACAAAGAAATGAGCAATTA 1318
QY 101 CGTAGTTGATAGCTGAACACAGGCTCATATAGATTAGAACAAACTGGCC---A 156
DB 1319 AGGAAATCTATTGCTGAACACAACTCATATTGCAAACTATTAAGATAGAGCCACAA 1378
QY 157 CAGTACTGAATGAGCCCTGGGAAAGCTTTTCTATCCAAAGAAAGTATGTGGACCCAG 216
DB 1379 CTAAGAGAAATTAACCTCGAGGAAGGGAATGTGCAAGAAATATCACAAAGACAGAA 1438
QY 217 ACCCTTACAGTCACTTAAGAAAGATTCAAAAAGCCTGCTGTGGCACTGGATGAAGCC 276
DB 1439 AACATGATATGCCCAATTAAGAGAGGTGCGCCAGCAGCCCTGCTGTGATGAAGCC 1498
QY 277 ATTTCTCAATCAATC-----AGTTTCATGACATAGATCAATGCTTGGAGAGC 327
DB 1499 GTGTCCAGTCCACACAGATTAACAGAGTTTCAATGATTAATTAAGTTCATGTTGAGCA 1558
QY 328 CTGAGACGATCGTGAACGTCTGAGGAGCAGCCCTTATCTGTGACAGAGTTGAGAG 387
DB 1559 CTGAGAGATCTTCTCTCGCTGCGCTGGATGCGCACCACTGATCCCTGTGAAGTAGACAG 1618
QY 388 ATCAAGAAACAGATCAGTGAATAATGAGATGTCTCAGTACATGAGAAAGCTACAGCG 447
DB 1619 ATCAGAGAGTGCATCAGTACATTAAGAGTCCACCGTGAAGCTAAGAAAACCTGACGCA 1678
QY 448 TTGATGAATCTTTAAACAGAGGAGAGAAATGATGCTAGATCTGGGGGAGATGAT 507
DB 1679 TCTTTGAGGCTTTGAAGCGCCGTGAGAGAGCTTATTTGACATCTCAAGAGAGACAC 1738
QY 508 AAAGACATATCTGCCAAAGCTGTTTCAGAGATAAGCTTGACCAATGGTTTTCATTGGAG 567
DB 1739 AAGATCTGGCTGCAAAAGAAATCCAGATTAATTTGATCAATGTGATTTCTTCTGGAG 1798
QY 568 AACATACACACATCTGTGGAAGAGAGGAGCCAAACTATCTGATGTGATGAGAGCTAGCA 627
DB 1799 GACATCAAAAGCTCGGGCTGAAGAAAGCAAAATCAAAATTTCTTGATGTCTTGAATTAGCA 1858
QY 628 GAAAGTTTCTGATGATCATATGATGATTAATGATTAACATTAAGATCACTCAAGATTC 687
DB 1859 GAGAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1918
QY 688 ATCCGGGAGCTGGAAGATCTGGAATTTGATCTTTCAGTATGATTAACACAGCAAGAGCA 747
DB 1919 GTCCATGATCTTGAAGAGCCAGGATGATCTTCTTCATCATCAACACAGAGTTGAGCT 1978
QY 748 GCAAGACCATTAAGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 807
DB 1979 GCTGAGACTATTAAGAGAGAGACAGATGTCTGATGATGATGATGATGATGATGATGAT 2038

QY 808 CTAGTTCTGAATCATATTCGCGCATGTGGGAGGAGCTGATTAACCATTTGCAAGAGAGT 867
DB 2039 CTTGAGAGAGATTTGATTTTGGCTGTGAGAGAACTGAGAGGCTGAAGTGAAGAGAGC 2098
QY 868 ATACATGATTAATTCAGCGATGGATTTCTTAATTAAGCTTGGAAAGACCGGATTTGA 927
DB 2099 ATTATGATGATGAATTAATGCT--TTGGAGAGACTTAACAAAACATGAAAGAGAGGCTAGA 2157
QY 928 CCAACTTGAAG-----AGCAATGACAGTCCCGCTCAGTACATGAGATGAGATGAGAGCTGAT 981
DB 2158 AAAACTTGAAGATCTTATGCAAGTCTGTGACATATGAGACACTTTCAGGCTATGTT 2217
QY 982 TTCTGGGTGATATTTGCAAGGTGTGTAAGTTGCTTCAATGCTCCAAATTTGAAACAGATC 1041
DB 2218 TGACTGGCTAGATTAACACTGTGATTAATCTGACCATGCCCCCTGTGGCAGTGC ACC 2276
QY 1042 TCGAACTGTCAAGCAGCAGATTTGATGCTAGAGCAATTTAAGTCTGAGGCTTATCAAC 1101
DB 2277 TCAATACTGTTAAAGATCAGTTAATGAATGAAGAGAGTTCAAGTAGAAGTTTACCAAC 2336
QY 1102 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
DB 2337 AGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2396
QY 1162 NAGAGATGACAAACACACTGTTCTAGACCCAAATATGANTGTATTTGATAT--GGCTA 1220
DB 2397 ATGAGAGGACAGAGACATTTATGAGAGACACTGACAGAACTCAACACCTCTGGAGAG 2456
QY 1221 GCCTGAGAGAGAGATCATCAACAGACAGCAATAAATGAGAGGAGTCTTATTAAGCCNTGG 1280
DB 2457 ACCTGGGTGAGAAATTTGGCCACCGACAGCAAACTGAAGAGGGGCTGTGTTGGCCCTTG 2516
QY 1281 GTCACTTCAACATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
DB 2517 GTCACTTCAACATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2576
QY 1339 TAACTGAGACAGAACTGTTGAGAGAGACCTTAAGCCATTAAGATTAATGAATTTGACCAAGC 1398
DB 2577 TGAATGCTCAGAGACCAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2636
QY 1399 ATCATGTGCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
DB 2637 ACATGTCCTTAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2696
QY 1459 CAGAAATGATCTTAATTTGATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
DB 2697 CTGGCAATGAGCTTCTTGAATCCAGTGTGAGATGATGAGAGAGAGAGAGAGAGAGAGAG 2756
QY 1519 TAGAGTTTAAATCAACGCTGGCAAAATGTTTGAAGAAACAGAAACAAAGAGAGAGAGAG 1578
DB 2757 TGAAGCCATGACCAATGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2816
QY 1579 AGCTGATGTGCTGCTGGCCGAGGAGCAAGAGGTTCAATGAGCAATTTGAGAGATTTGAGAGC 1638
DB 2817 AGCTTCAATGATCTGAG 2876
QY 1639 AGTGGCTGATGACAGAGAGGCTCATCTTTGAGATCTTAAACCCCTGGAGAGCTTACCGG 1698
DB 2877 TGAAGCTTACTAGATGAG 2936
QY 1699 AAACAGCAAG 1758
DB 2937 AAATGCTGAG 2996
QY 1759 AAGAAATATTAAGATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1818
DB 2997 AAGAGACTTATTAACACTTGAACAAGGACACATCATGCTTTTAAGCCGTGAGAGACT 3056
QY 1819 CTG 1821
DB 3057 CTG 3059

RESULT 13
AB007934 14491 bp mRNA linear PRI 28-AUG-2003
LOCUS Homo sapiens mRNA for KIAA0465 protein, partial cds.
DEFINITION AB007934
ACCESSION AB007934 GI:34328013
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Seki, N., Ohira, M., Nagase, T., Ishikawa, K., Miyajima, N.,
Nakajima, D., Nomura, N. and Ohara, O.
TITLE Characterization of cDNA clones in size-fractionated cDNA libraries
from human brain
JOURNAL DNA Res. 4 (5), 345-349 (1997)
MEDLINE 9811662
PUBMED 9455484
REFERENCE 2 (bases 1 to 14491)
AUTHORS Ohara, O.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1997) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
COMMENT On Aug 27, 2003 this sequence version replaced gi:20521060.
FEATURES
source location/Qualifiers
1..14491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="hg0112362"
/cissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/note="This sequence was obtained by subcloning of the DNA
fragments derived from three cDNA clones (1 - 3992 was
derived from hb07943, 3993 - 8100 was derived from f10348
and 8101 - 14491 was derived from pf00419)-This sequence
was replaced that of hg01123 cDNA as a representative cDNA
sequence for KIAA0465."
1..14491
/gene="KIAA0465"
<1..13303
/gene="KIAA0465"
/note="Start codon is not identified"
/product="KIAA0465 protein"
/protein_id="BAA32310.3"
/db_xref="GI:34328014"
/translation="KNTISVKNVCYROIETICRNDCEVLNRSQRTKMKYISPTGN
EAMVPSVCLIPPNKDAIEMASRVQSGYKWMALHQLHNTKSLISMYLAKDLDL
VOTWLEKRSAPGECIOIMKQLAHYEDFDLSDSDVLFVADRLRLSEEVACAKA
RQHLKMSMENEDKEETVAKMYISELKNIRLREYEVQVVKRIQSLASRRDRDMQ
DNALRIABOEHTQEDLOQLRSLDAVSMKCDSPSSSVPTLRSELNLVEKMD
HYGLSTVYLKTKTVDIVIRSIQDAELLVKGEIRLSQEBVVIALDLSEAHWTFMD
HMLSDVKNSVSVLDEIRAKKVAABMSRLTPERNLDLEVOEKSGLOERHWRV
IAQLEIRQSELSIQVGLDYRACHSTLIKWIETTAQGBMKPGGAEDSRVLSBUS
OQTALPABIRNOTKLDQCKTSQYTSIVKDELMTYKKAVEREQSGPKGRRLML
SSSDATVQEFMDLRTYALVTLTTOHVXYISALRLREBEKRVSEKQEHVEKQIF
LKGWSTLARNTOGKATSEETKSTDEKALIEQVLSLEELATRKQVSEATITQIF
LNHGKLSSEKKEKQISEQLNLNKAHYHLDGDSANQLQOLQSLAHOTBKTLQKQ
NTCHOOEDLCSVGOABERAIAGHGRRTQODLSALOKNOSDKLDODIQRATSEFA
TYVKDIEGFMERKQTKLSPRELALBEKLIHQAEQYEAQLEBTRVAKQELBEVTSAL
QOSTEKSAKAEIENKKKI DALDWNTVSGSGQLNLNPEMQLSGASLEKGLD
TIDGYVNAQAPKPKLQCEMKMKARHQLSOQNLILATQSOAQLDGHNLPTPEE
OQMLQKGLKEQYSTLSAQSBAEKQVLODELOKFTLQDHKESEKSEKLE
NNHKGSSPETLPSLKROGSPSEDIYSHKGRFPYITIGOKYIMDENSEPKSEKSE
IGNLVKDKLDAETERYTALSKCTRLGSHLNNLTIGYHOPNSADSLQAMQACENY
EKLSPTVASDPROVLOEQATTQLOLEBIAEHQVPEKQIKVARDIMEIIEGEAPDHR
HVOETDSTLSHFQSTLSYSLAERSLLOKRAIQSOSVQSLSELSLQISIEVQNLGK
QVSSLSGVIQEBALATMKLKDPIAROKSSLEATREMTVRFMETADSTTAAYVLOGKLA

EVSORPEQLCLOQEKESLKLPLPOAENFEHLSGKLQOEMENKSHMLASGNPDODI
THFPOQIEALNEMEDQDENTLDELHLYLFTKELSCGALDLCQHDQVQMLRKPTFLQ
KTYFEREKDASSCOEOLDEFRKLVTRFOKMLKETSISIPPTESMAKAELEKQIEHLK
SLIDPMASGKTVETINCKTDLNIMETITPDSQKDLNLEIQCDDMDVNYKVEYKL
GGYTLHEROSLQAILNRMERHKEANSVITQMLBSKEVTKSHDMSPTKTEYVAKQ
ESNKAPLAELEQNSPKIIOVKALAGLAVTTPNSDAENWKIIOEILNRWEATVET
VARORLREESASHLACFOAASQLRPMWMEKELMMGVGLPISIDPMNLNAQOQVFM
LKEPEARQOEQLNEAOGILITGPVDVLSISQVKELOSINQKVELTQORLKNSSS
QIDQAVKSTORYOELLQDLSKRYAVAGLVSQALSTQPEAVKQOOLESEIRLBLE
QDHEVKEAQTCDELSVIGBOYIKDELKRLFTYALPLQGLIEDLAADRIRNLQAL
ASTQOQOMFDELRTLDKQSOQANCTISAKLERLOSQLOENEFQSGLNQSSY
EVIIVAGESLILSVPEGEKRTLQNVLEKSHWELSKTDRORSRLDCQKAKY
QMHVEDLVWIDCKAKMSLVTTLDPVLEBSLSKMSKMLVEVERBRLLELINSAA
DILINSSEADBDGIDDEKAGINMDAVTEIEQAKTSQLEETQORLEFQESKATK
KVEGAKHOEITFDALGSOASNKNEKLAQOEVOLAPPOVDYLRFPFGVYEDAP
GSDASQLHQAEVAAOEFLEVQORVNSGCMVMENTLQEGORHCRMEFESQALDD
ELDMGATIRDTDSLOSQIEDVRLPLNTKTHVLKDIIEASBAGRHLDEGDLILGL
KRELEALNKQCGKLTERRKQOEULETLGRVEDFYRKLGKLNDAITTALEBVALQW
VGEVEVILNQQLADPFMFQOEQVDPLOMKLOVNGIGQGLIQSAGDCDVOGHEHME
EINARWNTLNKKYAOIRIAQJALHCGFKODLEPLLSMLADTEELINOKPSPAEY
KYVKAQIOEQLORLDRKATVMDLQAEGRILQSAELADREKITTGOLSELSWT
ELLSKAAQKOLEDLIVAKOPHETABEISPLSTYTEKKLANSBVGOTAKIOQOI
IHKALNEBTVNRKNVDOAIKNGQLKOTIGEEVULIOEKLDGIKTRVADITVTS
KALRTLQOROLATKFOSTYBELTGMRLREVEBELATVSDNGQSPTEQOIPQOQRKELK
KEVMEHRLVLDIVNEVSRLBELVPMRAREBGDLVSDNEQYKLVSDITGRVDEID
MAIORSQOEAADALAVATKRLMALGPIRLQODOTKALQVKAFSIDIIRK
DAMDLEFHSRSEIFGTGGEOKTVIOEKTESLIOQYEAISLNSERYALEROVAVN
OFWEYEEELSPMIEETRALIOLPSAIDHBOLOQOQEMEROLRESIAHKHPIDKL
KIGPOLKELNPEGEVMEBETKQAEKMAIKRECISSDKSAVLELEKLDPAVSOSTOTTERDK
IIPMLTEENLSSRLRMPPLIAEVDKIRCEISDNKSAVLELEKLDPAVSOSTOTTERDK
LIGRSGADKDLAAKEIOPDKLOVWFPMEDIDKARAEERIKELTDVLELEKFGYDAA
LITTIKQDQIDVHDESIGIDPSIIKOQYEALETKIEEDGHELEFRIIGADLIF
ACGETEKEPVSIDEMNNAMENKNTKTEVREYKLEEDANAQVAYODDTLOAMEMDTN
TVYKLCMPVSGTDLNTVYDQNLNEMKEFEVRYOQOILEMKNHOSELMALKATDSTD
RQIIEPLTELHLENBCEKTAHROHKLBSGLLAGQOHALBEMLSCVLTPTTEILD
ACRPIISGDKVELELAKHNLKNDVLAQVAFETVNRKGNELSESDADNLSR
LBAQMCSVLOKTEBERQOOSTLOQAGHVESEIDELLETITRESQDLSASPKPG
LBEPTAREODITMEIYLSOLKAEETYNQLDGRMLLSESDSGSEKTSQSVALLQD
KMHVSSKMEERKSLEBMLNATERONSLOPFIWMLTABSOLINASPISLILNLV
SOIEBHRYANVANHRODITILDQNGOKRLSOKQOVVILKNLVSVQSRREKVO
RSIEGRSLDDARKKQOPEHAKWKILDLBEAESHLSDELISNDPDKIKQLSVRQ
BFQKTLGQKQVYDITTRTGRALKERTLLPEBSOKLNDLFGVDRKIDPQVCGVSRQ
HKLBEALLESQFMDALQALVDMLYKVELELDEDPVHGDLDLVNMLAMHAKYFQEL
GKRTGVQVLYKRSQOELIENSDDFTWVGQLOEISTRDYDCKLSVSKOSRLQALK
QAEVFRDPTAHMLLMLSEAOITLRFRGLPDDTLALOSLIDHMKMKKVEBERVUN
SAVAMGEVILAVCHPDCTTTTKMTITIRAREBVLITAKQHQKLETLRLSELVANAE
LLEELIAMIQMAETLILQDQEPITQNDIRVALIAEHOTFMEBETRKOPDVDRVXT
YRKRIEPTHAPFIEKRSRSGKSLSQPPRPMPTIISQSEANPRINQJASMQOQWL
LALEROKLNDALDLLELEKEFANFDPVWRKYRMWNNKRSRVMDFRRRIDKQDQ
KITROBEFIDILASKEPTTKLEMTAVADIFDSGQGYIDYEFVALAHNKQAVRETT
DADKTEDVTRGVACCKAKRPOVQOIGENKXRPDQSOLRLVRLILSTRVMMRVGSGW
NALDEFVLVNDCCRAGRNIILREKFTILPEBASQGMTFRBGRGRSKRSRPAASPTR
SSSASQSNHSCITMSPSSPATPASGTVIPSSGSKLKRPTLPFHSSRTSLADTSSS
SPASGAKETNRADPKPSARPSGRSAGRSASRSRGSADBDFLLETQSCASDTSE
SSAAGQGSNRGRGLNPKSKIPTMSKTTTASPRTPGPKR"

ORIGIN

Query Match	32.3%;	Score 655.4;	DB 9;	Length 14491;
Best Local Similarity	62.9%;	Pred. No. 3.9e-148;		
Matches 1134;	Conservative	0;	Mismatches 645;	Indels 24;
			Gaps	7;
QY	41	CCGCGCCGAGCCCTTGAATATATCTCTAAGACGACGACAGAGAAATGAGTGGCAACTG	100	
DB	8555	CCCTCTCCAGCCATTCATATATGACAGCTCAGGACGACACAAAGAGAAATGAGGCAATTA	8614	
QY	101	CCTGAGTTGATAGCTGGAACACAAAGCCTCATATAGATTAGATGAACAAATGGCC----	156	
DB	8615	AAGGAATCATATTGCTGGAACAAACCTCATATTGACAAACTCTAATAAGATTAGGCCACAA	8674	
QY	157	CAGTACTGATAGAGCCCTGGGGAAGGCTTTTCTATCAAGAGAAAGTATGGCAGCCGAC	216	
DB	8675	CTTAAAGGAATTAAACCTTGAGAGAGGAGGAATAGTGGAAGAAATAATCCAGAAAGCAGAA	8734	

217 ACCCTTACAGTCACATTAAGAAGATGTCAAAAAGCGTGTGTCAGTCAGTGAAGCC 276
8735 AACATGATGCGCAAAATTAAGAGAGAGTGGCCGACGACCCCTGGCTTGTGATGAAGC 8794
277 ATTTCATCACTC-----AGTTCATGACAAATGATGATCAGATCCTTGAGAGC 327
8795 GTGTCGCCAGTCCACACAGATTAACAGAGTTTCATGATTAATAATTGAGCCTATGTTGAGACA 8854
328 CTGGAAGCCATCGTGGAAAGCTGTAGGAGGAGGAGGAAATGATGCTGAGTCTGGGAGAG 387
8855 CTGGAAGAAATCTTCTCTGCGCTGCTGATGACACACATGATCCTGCTGAAGTACAG 8914
388 ATCAAGAACAGATCAGTGAATAATTAAGATGTGTGATGACATGGAAGAGTACAGCG 447
8915 ATCAGAGGTGATCAGTGAACATTAAGAGTCCACCGTGGAGCTGAGAAAACCTGACGCA 8974
448 TTGATGAAGACTCTTAACACAGAGGAGGAGGAAATGATGCTGAGTCTGGGAGAGCTGAT 507
8975 TCTTTGAGGCTTGAAGCGCGTGGAGAGGAGCTTATTGAGCATCTCAGGAGACAGAC 9034
508 AAAGACATATCTGCCAAAGCTGTTCAGAGTAAAGCTTGAACAAATGCTTTTCAATTGGAG 567
9035 AAGATCTGCTGCAAAAAGAAATCCAGATTAATTTGATCAAAATGATTTCTTCTGGAG 9094
568 AACATACACACACTGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 627
9095 GACATCAAGAGCTCGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9154
628 GAAAGTTCTGTGTGATCAGATGTCATTTGATTAAGTAAAGTAAAGTAAAGTAAAGTAA 687
9155 GAGAAAGTTCTGTGTGATCAGATGTCATTTGATTAAGTAAAGTAAAGTAAAGTAAAGTAA 9214
688 ATCCGGAGCTGGAAGATCTGGAATTTATCTTCAAGTAAAGTAAAGTAAAGTAAAGTAA 747
9215 GTCCATGATCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9274
748 GCAGAGACCATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 807
9275 GCTGAGACTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9334
808 CTAGGTTCTGAATCTATTCGCGCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
9335 CTTGAGAGAGATTTGATTTTGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9394
868 ATACATGATTTAAATTCAGGAGTGTCTTAATTAAGCTTGAAGAGGAGGAGGAGGAGGAG 927
9395 ATTATGATGATGAATTAATGC-TTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9453
928 CCAACTTGAG-----AGCAATGACAGTGGCGGTGATGATGATGATGATGATGATGATGAT 981
9454 AAAACCTTGAGAGTGTGATGCAAGCTGTGCAAGTATGAGGAGGAGGAGGAGGAGGAGGAG 9513
982 TTTCTGGTATGATATTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1041
9514 TGACTGGTATGATTAACATGATTAATTAATCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 9572
1042 TCGAAAGCTGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
9573 TCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 9632
1102 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
9633 AGCAATTTGAGATGAGAGGCTTAATCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9692
1162 NAGAGAGTGAACAAACACATGTTCTAGACCCAAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220
9693 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9752
1221 GCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1280
9753 ACCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9812
1281 GTCAAGTTCAACATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338

9813 GTCAAGTTCAACATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 9872
1339 TAAGTGAACGAAACCTGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1398
9873 TAGATGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9932
1399 ATCATGCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
9933 ACATGCTCTTAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9992
1459 CAGGAAATGATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1518
9993 CTGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10052
1519 TAGAGGTTTAAATCAACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1578
10053 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10112
1579 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
10113 AGCTTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10172
1639 AGTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1698
10173 TGAAGTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10232
1699 AAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1758
10233 AAAGTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10292
1759 AAGAAATCTAATTAAGTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1818
10293 AAGAGACTTAAATCAACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 10352
1819 CTG 1821
10353 CTG 10355

RESULT 14
AB029290 17518 bp mRNA linear PRI 03-SEP-1999
LOCUS
DEFINITION Homo sapiens mRNA for actin binding protein ABP620, complete cds.
ACCESSION AB029290.1 GI:5821433
VERSION actin binding protein ABP620.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 17518)
AUTHORS Matsuda,S. and Okuda,T.
TITLE Actin binding protein ABP620
JOURNAL Published Only in Database (1999)
REFERENCE
2 (bases 1 to 17518)
AUTHORS Matsuda,S. and Okuda,T.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-1999) Satoru Matsuda, Nagoya University School of
Medicine, Molecular Pathogenesis, 65 Tsurumai-cho Showa-ku, Nagoya
466-8550, Japan (E-mail:smatsuda@bsu.ru.med.nagoya-u.ac.jp,
Tel:81-52-744-2463, Fax:81-52-744-2464)
FEATURES
source
1..17518
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HePG2"
1..17518
/gene="abp620"
52..16344
/gene="abp620"
/codon_start=1

Oy 688 ATCCGAGACCTGGAAGATCTGGAATTGATCTTCACTAGTAAACAAGAGAAAGCA 747
 Db 12256 GTCCATGCTTGGAAAGCCAGGCACTGATCTTCCATCATCAACCAAGGTTAACT 12315
 Oy 748 GCAGAGACCATTAAGGAGAAATAGATGACTACAGAGAGAGCTGATATATTAAC 807.
 Db 12316 GCTGAGACTATTAAGGAGAGAGAGATGCTGCAATGAAGAGCTGATATTAAC 12375
 Oy 808 CTAGGTTGAACTCATCTGCGCATGTGGAGACCTGATTAACCCATTGTCAAGAAAGT 867
 Db 12376 CTGGAGAGATTTGATTTTGGCCGTGGAAGAACTGAGAAAGCCGTAAGAGAAAGC 12435
 Oy 868 ATACATAGTTAAATTCAGGAGATGGGATTTCTTAATTAAGCTGGAAGCCGATGA 927
 Db 12436 ATTGATGATGATTAATATGCTTGGAGAACTTAAACAAACATGGAAGAGAGCTAGA 12494
 Oy 928 CCAACTTGAG-----AGCAATGACAGCTGCCGTCACTACAGATGAGACTGACAGCGGTAT 981
 Db 12495 AAACTTGAAGATGCTATGCAAGCTGCTGTGCAAGTATGAGACACTCTTCAGGCTATGTT 12554
 Oy 982 TTTCTGGTATATTTGCAAGGTGTAAGTTCGCTTCAATGCTCCATTTGGAACAGATC 1041
 Db 12555 TGACTGGCTAGATTAACACTGTGATTAACCTGACACATGCCCTGTTGGCACTG-ACC 12613
 Oy 1042 TCGAACTGTCAACAGAGAGATTAATGCTAGAGCAATTAAGCTGAGCCCTATCAAC 1101
 Db 12614 TCAATACCTGTTAAAGATAGTTAAATGAAGAGAGCTTCAAGTAAAGTTTACCAAC 12673
 Oy 1102 AGCAGATAGAGTGTGACAGCTGACTCATACGACAGACCTTTTGTCTAAGAAAGTACAG 1161
 Db 12674 AGCAATTTGATGATGAGAGAGCTTATATCAACAGGAGTGAATGATGTTAAAGAAAGTACTG 12733
 Oy 1162 NAGAGAGTGAACACACACTGTTCTAGACCCAAATGAGCTGATATTTGAT- GAGTA 1220
 Db 12734 ATGAGACGAGAGAGACATTTATACGAGAACCACTGACAGAACTCAACACCTGAGGAGA 12793
 Oy 1221 GCCGTGAGNAGAGATCATCAACAGACAGCACTAACTGAGAGGTCTTATTAAGCCTG 1280
 Db 12794 ACCGTGGTGAAGAAATGCGCCACGACAGCAAACTGAAGAGGCGCTGTTGGCCCTTG 12853
 Oy 1281 GTCACTTCCAACTATGCTGTATGAGCTCCGTCAT--GCTGACACACAGAGGCTTGC 1338
 Db 12854 GTCACTTCCAGACTGCTTAAAGAGAACTAATGATGCTGATCTATACCAAGAGTTGT 12913
 Oy 1339 TAACTGAGACAGAACTGTTGAGAGAGACCTTAAAGCCATTGAATTGAATTTGCCAAGC 1398
 Db 12914 TAGATGCTCAAGACCAATTAAGTGAAGACCAAAAGTCAATTGAAGTTCGCAAGC 12973
 Oy 1399 ATCATGTGCTCCAAATATGATATTAAGCCATCACTGACAGTGAAGCCGTTAATTAAG 1458
 Db 12974 ACCATGTCTTAAATATGATATTTTGGCTCATCAAGCAGTGAAGCAAGTCAACAAAG 13033
 Oy 1459 CAGAAATATGATCTAATTAATCAAGTGAAGAGAAAGAGCAAGCACTTCAGAAACAGC 1518
 Db 13034 CTGGCATGAGACTTCTGAATCCAGTGTGAGATGATGCGACAGCTTAAAGAGCCGTT 13093
 Oy 1519 TAGAGGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAAACAAGCAAAAGAGAGCAGC 1578
 Db 13094 TGAAGAGCATGAACCAATGCTGGAGAGTCAAGTGTTAACAAGAAACAGAGAGAGGAGCAGC 13153
 Oy 1579 AGCTGATGATGCTTGGCCAGAGCCAAAGGTTTCATGCGGAATTTGAGATTTGCAGC 1638
 Db 13154 AGCTTCACTCAACTCTGAGAGAGGCCAGGCTTCCACAGGAATTTAAATTTCCCT 13213
 Oy 1639 AGTGGCTGACTGACAGGAGGCTCATGTTGGCATCTTAAACCGGTGAGAGTTTACCGG 1698
 Db 13214 TGAACCTTCTAGATGAGAGAGCCAGCTTTCTGATCTTAAGCCACAGAGAGAGATTCTTG 13273
 Oy 1699 AAACAGCCAGAGAGAGCTTAAATGCTCATATGAGAGTCTGTGCTGCTTTGAGAGTAAAG 1758
 Db 13274 AAATGCTAGAGGAAACAGCTGATACATATGAGAACTTATTTCCAGCTGAAAGCAGAG 13333

Oy 1759 AAGAACTATATAGAGTCTGATGAGAGAGCCAGACAGTCTGCTCAAGATGCCCAAT 1818
 Db 13334 AAGAGACTTATTAACACTTACTGACAGAGGACAGACTCATGCTTAAAGCCGTGACACT 13393
 Oy 1819 CTG 1821
 Db 13394 CTG 13396

RESULT 15

AF141968

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

CDS

FEATURES

source

1. 17734
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p33-p34.2"
 965..17086
 /note="cytoskeletal giant protein; actin-binding protein"
 /codon_start=1
 /product="trabeculin-alpha"
 /protein_id="AA06360.1"
 /db_xref="GI:6273778"
 /translation="MPPVLMAGIPGRDVSLQPLPQFKQCTSSASHVAAYADRDVY
 OKTFTKWKVHKHMKVKKHINDLYEDLRDGHNLISLEYSGLIPREKGMFRHRLQ
 NQGLADPLKQROVKLVNIRNDITDQNPGLTGLITVTLHPDISIYISGESGMS
 AKELULMTQKVTAGTIGTICNTSSCSWGSKFNALHHRPDLVMEKQIQSNRE
 NLEQAFVAVRLGVTRLDAEDVDVDPDKSVITTVSSLYDAVPKVEGEGISATE
 VDSRWQYQSVRLSPWIKQHTILMSDKTFPPQVPLKALYNQYHFKETELAKSR
 EKGRIELVLYLWVIEFGRIKLPGQHNPDEBEMKLIEMLEKSRPAVERLE
 LLQIANKINGLNCBEXLTKLNTLOADAHLNBSGQPOCESDVTWYIOEGGLR
 QLOVDLQILEDENYQLEELAFVMPRLDQDLVTLRECTNLYRGHFTSLVPSPL
 TTHLKAEPILTKATHSSSTFWKPMTRAEVLVSSSEDEGNLRFVYELSWBEMQ
 KLERAEWGNLPVELQLLETOQHILHSVEBELVASVKARLYEGMSQNPFTSVETG
 KLITQYCKLKETSPFRMRHLOSHPKFSRTAEILWNEBEELALDWSNNNSIA
 KRVFSELMELEKQVPRSLQDTALSLLENHPAQVTEAVSAQAQSLQMKQLC
 LCVQHYKENTAVQFSDARELESFRLNODSIKRYSCDHNTLSRLLEDLQDSMD
 EKQQLQSKSSVSLVGRSKTIYQKRSDDHLYKTIYKAVQDVQOIEITCKNE
 CVLEDNSQRTKMKVISTFGNEMAVPSCFLLPPNKQALIMASKEVDSYQKVALMHQ
 LHVNTKSLISWNYLRKDLVQTNMEKLSAAGECHQIMKQLQAHVEDFLQDSRS
 LVPSVADRLEEBEVEKCARFQHLNKMSENEDLETVAKVYISLKNILRLREBYO
 RVVRIQSLASRTDRDAQODNALRIEGRHEDQLQSLDADAVMKDCFLHQP
 SSSVPTLRSELNVLVKMDHVGSLTVYLNKLTQVIVIRSIQDAELLVKGYEIS
 QERVVLADISALFAHSGTLRHLMSDVQKSVSVSDEIARAKVVAEQMSRLPREN
 LDLERYQKSSQLOERHVRVIAQLEIQLQSELESTIQVLDGTRFACHGTLIKRIETTLQ
 QEMKPGQADRSQLESLQQTALFELIERNQTKLQDCQKFSQOQSYITVYKDELQAM
 TYKAFVDSQKSPKRRRLMSSDAIQLQEPMDIRYRTALVTLTQHVAKYISDLARL
 EEBKRVVEEKQSEVEKVELLQVSTLQNTQKATSEKSTEDIEKALILEQVLS
 EELTTKEQVSAIKTSQIFLAGHGLSEKQKQISEQLNALNKAHYHDCDSANLQ
 QQLQSLAHTQETLQKQNTCHQQLLEDLCSWQAEKRLAAGHQTTOQDLSALQK

Db 13419 AGCAATTGATGAGAGGCTTAATCAGCAGGTGAAGTGTAAAGAAAGCTACTG 13478
QY 1162 MAGAGTGAACAAACACACTGTTTAGACCCAAAAATGAGTGTATATGATAT-GGCTA 1220
Db 13479 ATGAGAGGAGACAGATTAATACGAGAACCTGACAGAACTCAAAACACTCTGGAGA 13538
QY 1221 GCCTGAGAGAGATATCAACAGACAGATAAAGTGAAGGCTCTTATTTAGCCNTGG 1280
Db 13539 ACCTGGGTGAGAAATTTGCCACAGACAGCAAACTAGAGGGGCTCTGTTGGCCCTTG 13598
QY 1281 GTCAATTCCAAATGAGTGTATGAGTCTGAT--GCTGACACACACGAGGGCTTGC 1338
Db 13599 GTCAATTCCAGCATGCTTATAGAGAACTAATGAGTTGGCTGATCATACGAGAGTTGT 13658
QY 1339 TAACTGAGCAGAAACCTGTTGAGAGAACCTTAAAGCATTGAATTGAATTGACCAAGC 1338
Db 13659 TAGATGCTCAGAGACCAATAGTGAAGAGCCAAAGTCATTGAATTGAGCTCGCAAGC 13718
QY 1399 ATCATGTGCTCCAAATGATGTATTTAGCCCATCAAGTCCACAGTGAAGCCGTTAATAAG 1458
Db 13719 ACCATGTCTTAAAAATGATGTTTGGCTCATCAAGCCACAGTGAAGCAATCAACAAAG 13778
QY 1459 CAGGAAATGATCTAATTGAATCAAGTGCAGAGAGAAAGCAAGCACTTCAGAACAAAGC 1518
Db 13779 CTGGCAATGAGCTTCTGAATCCAGTGTGAGATGATGCCAGAGCTTAAGAGCGCTT 13838
QY 1519 TAGAGTTTAAATCAACGCTGGCAAAATGTTTGGAAAAACAGAACAAAGAGCAGC 1578
Db 13839 TGGAAAGCCATGAAACCAATGCTGGAGAGTCAGTGTTAACAGAAAAACAGAGAGAGGAGCAGC 13898
QY 1579 AGCTGSAATGATGCTGCGCCAGGCAAGGGTTCCATGGCGAAATTGAGGATTTGCAGC 1638
Db 13899 AGCTTCACTCACTCTGACAGAGGCCAGGCTTCCACAGTGAATTAAGATTTCTCT 13958
QY 1639 AGTGGTGAAGTGAACAGGAGCGTATCTGTTGGCATCTAAACCGCTGGAGGTTTACCGG 1698
Db 13959 TGGAACTTACTAGATGAGAGCCAGCTTCTGCACTTAAGCCACAGAGAGACTTCCTG 14018
QY 1699 AAACAGCCAGAGAGAGCTTAATGTCATATGAGAAATGTTGCTGCTTTGAAGCTAAAG 1758
Db 14019 AAATGCTAGAGGAAACAGCTGATACACATATGAACTCTATTTCCAGCTGAAGCCAGG 14078
QY 1759 AAGAAACATATAAGTCTGATGAGAAAGGCGAGAGATGCTTGCAAGATGCCAAAT 1818
Db 14079 AAGAGACTTAATCACTACTGACAGGCGAGCTCATGCTTTAAGCCGTGACGACT 14138
QY 1819 CTG 1821
Db 14139 CTG 14141

Search completed: January 12, 2005, 11:10:12
Job time : 8369 secs

THIS PAGE BLANK (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 06:55:34 ; Search time 5752 Seconds
(without alignments)
12866.672 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ttgcacatgcgcacgaaga.....tngagagctcccaacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633.6	31.2	1053	4	BG289408 602381438
2	623.6	30.7	676	5	BK472284 DKF2p686E
3	618.4	30.4	815	7	CK000929 AGENCOURT
4	561.6	27.7	764	7	CR549848 DKF2p469K
5	526.8	25.9	752	7	CK781134 UI-M-GVO-
6	486.4	23.9	800	5	BU701246 UI-M-EXO-
7	481	23.7	643	5	BK470591 DKF2p686M
8	478.2	23.5	589	7	CO700994 DG32-191b
9	469	23.1	469	6	CB164054 K-EST0225
10	468	23.0	826	5	BU139957 603134131
11	455.8	22.4	724	6	CB525847 UI-M-FY0-
12	449.8	22.1	799	6	CB722030 UI-M-GH0-
13	441.6	21.7	836	5	BU217687 603759087
14	440	21.7	793	5	BU368661 603597561
15	437.6	21.5	849	5	BQ215351 AGENCOURT
16	432.2	20.3	622	6	CA751356 UI-M-FOO-
17	413	20.3	511	1	AI889212 t250d03.x
18	406.2	20.0	749	7	CP743131 UI-M-GVO-
19	404.8	19.9	641	6	CB520499 UI-M-GIO-
20	403.2	19.9	748	6	CB520969 UI-M-GIO-
21	400.8	19.8	737	6	CB520605 UI-M-GIO-
22	400.8	19.7	703	4	BI731712 60353775
23	396.4	19.5	571	1	AI526522 u39602.Y
24	395.2	19.5	648	4	BM388627 UI-R-D20-

c	25	393	19.4	512	2	BE080914
	26	389	19.2	723	5	BU305460
	27	386	19.0	778	5	BU748872
	28	367.2	18.1	665	6	CD804239
	29	366.6	18.1	665	6	CD804438
	30	366.6	18.1	665	6	CD804591
	31	361	17.8	740	6	CB521821
c	32	358	17.6	361	4	BG949751
	33	353.2	17.4	486	4	BI345034
	34	345.6	17.0	490	6	CB548311
	35	340.2	16.8	719	5	BU477650
	36	335	16.5	805	6	CF284483
	37	333.8	16.4	446	2	BE871902
	38	328	16.1	331	2	BE003231
	39	322.8	15.9	533	2	AM475236
	40	321	15.8	427	6	CB794802
c	41	320.6	15.8	470	2	BE081412
	42	319.2	15.7	474	5	BK475908
	43	313.6	15.4	693	7	CN420731
	44	312.6	15.4	656	7	CR536640
	45	309.8	15.3	425	7	CK393268

ALIGNMENTS

RESULT 1
BG289408 1053 bp mRNA linear EST 21-FEB-2001
LOCUS 602381438F1 NIH_MGC_93 Homo sapiens CDNA clone IMAGE:4499097 5',
DEFINITION mRNA sequence.
VERSION BG289408
KEYWORDS BG289408.1 GI:13045223
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10362 row: 1 column: 10
High quality sequence stop: 730.

FEATURES

source
1..1053
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4499097"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_93"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

ORIGIN

Query Match 31.2%; Score 633.6; DB 4; Length 1053;
Best Local Similarity 95.7%; Pred. No. 9.8e-143;
Matches 748; Conservative 0; Mismatches 24; Indels 10; Gaps 9;

QY 181 GGCCTTTCTATCCAGAGAAAGTATGTGGACCGGACCCCTTTACAGTCACTTAAAGAA 240
DB 1 GGCCTTTCTATCCAGAGAAAGTATGTGGACCGGACCCCTTTACAGTCACTTAAAGAA 60
QY 241 GAGTCAAAAAGGCTGC-TGTGGCAGCTGATGAAAGCATT-TCTCAATCAA-CTCAGTTC 297
DB 61 GAGTCAAAAAGGCTGC-TGTGGCAGCTGATGAAAGCATT-TCTCAATCAA-CTCAGTTC 120
QY 298 CATGACAAAGATAGATCAGA-TCCCTTGAGAGCC-TGGAACGATCGTGAAGCTCTGAGGC 355
DB 121 CATGACAAAGATAGATCAGATTCTCTGAGAGCCGCTGGAACGATCGTGAAGCTCTGAGGC 180
QY 356 AGCCACCTCTATCTCTGAGAGGTTGAGAAATGACAGATCAGATCAGTGAAGTAAAG 414
DB 181 AGCCACCTCTATCTCTGAGAGGTTGAGAAATGACAGATCAGATCAGTGAAGTAAAG 240
QY 415 AATGTGCACTAGACATGAGAAAGCTACACCCCTGTGATGAAACTCTTAAACAGAGGGA 474
DB 241 AATGTGCACTAGACATGAGAAAGCTACACCCCTGTGATGAAACTCTTAAACAGAGGGA 300
QY 475 GAGAAATGATGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCGCAAGCTGTTGAG 534
DB 301 GAGAAATGATGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCGCAAGCTGTTGAG 360
QY 535 GATAGCTTGACCAATGATGTTTCATTGAGAGAACTACACACACTGCTGAGAGAGAG 594
DB 361 GATAGCTTGACCAATGATGTTTCATTGAGAGAACTACACACACTGCTGAGAGAGAG 420
QY 595 GAGAGCAAACTACTGATGATGAGAGCTGACAGAAAGTTCTGCTGATCACTGTCA 654
DB 421 GAGAGCAAACTACTGATGATGAGAGCTGACAGAAAGTTCTGCTGATCACTGTCA 480
QY 655 TTGATAGTTACATTAAGATCTCAAGATTTTCATCCGGAGCTGGAAGATCTGGAAT 724
DB 481 TTGATAGTTACATTAAGATCTCAAGATTTTCATCCGGAGCTGGAAGATCTGGAAT 540
QY 715 GATCTTCACTAGTAAACCAACAGCAAGCAAGCAAGCACTAAAGGAGAAATAGAT 774
DB 541 GATCTTCACTAGTAAACCAACAGCAAGCAAGCAAGCACTAAAGGAGAAATAGAT 600
QY 775 GATCTCAAGAGAGAGCTGATTAAGTATTAACCTAGTCTGAACTCATTTGGGCAAT 834
DB 601 GATCTCAAGAGAGAGCTGATTAAGTATTAACCTAGTCTGAACTCATTTGGGCAAT 660
QY 835 GGGAGAGCTGAT-AAACCATTTGTCAAGAGATATCATGATTAATTCAGCGATGG 893
DB 661 GGGAGAGCTGAT-AAACCATTTGTCAAGAGATATCATGATTAATTCAGCGATGG 718
QY 894 ATTCTTAATTAAGCTTGGAAAGACCG-GATTGACCACTTGAGAGCAATGCAAGCTGCC 952
DB 719 GATCTTAATTAAGCTTGGAAAGACCGCAATTTGGCCAAATGGAAGAAAGCAAGAGC 778
QY 953 GT 954
DB 779 GT 780

RESULT 2
BX472284 676 bp mRNA linear EST 04-SEP-2003
LOCUS BX472284
DEFINITION DKFZ686E12132_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone
ACCESSION BX472284
VERSION BX472284.1 GI:3166552
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 676)
AUTHORS Bahr,A., Lauber,J., Mewes,H.W., Weill,B., Amidi,C., Oeinger,A.,
Fodor,G., Han,W. and Wiemann,S.
TITLE EST (Bahr,A., Lauber,J., Mewes,H.W., Weill,B., et al.)

JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sequence available.
This clone (DKFZ686E12132) is available at the RZPD in Berlin.
Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ686E12132"
/dev_stage="adult"
/lab_host="DH10B"
/clone_11b="686 (synonym: h1cc3)"
/note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIb;
cDNA-collection"

ORIGIN
Query Match 30.7%; Score 623.6; DB 5; Length 676;
Best Local Similarity 97.6%; Pred. No. 2,4e-140;
Matches 652; Conservative 0; Mismatches 13; Indels 3; Gaps 2;
QY 1139 GCTTTGCTAAAGAAAGTAAACAGAGAGTACAAACACTGTTCTAGACCAAAAT 1198
DB 9 GCTTTGCTAAAGAAAGTAAACAGAGAGTACAAACACTGTTCTAGACCAAAAT 68
QY 1199 GGAAGTGAATTTGATAT-GGTACCTGAGAGAGATCATCAACAGACATTAAT 1257
DB 69 GGAAGTGAATTTGATATGAGATAGCTGAGAGAGAGATCATCAACAGACATTAAT 128
QY 1258 GGAAGTGAATTTGATAGCCTGAGTACCAACATGCGCTGATGAGACTCTG-CAAT 1315
DB 129 GGAAGTGAATTTGATAGCCTGAGTACCAACATGCGCTGATGAGACTCTGAGACT 188
QY 1316 GCTGACACACACGAGGCTTCTAAGTACAGAGAACTGTTGAGAGAGACCTTAAAG 1375
DB 189 GCTGACACACACGAGGCTTCTAAGTACAGAGAACTGTTGAGAGAGACCTTAAAG 248
QY 1376 CATGAAATTTGAATCTTCCAGATCATGTGCTCCAAATGATGATTAAGCCATCAGTC 1435
DB 249 CATGAAATTTGAATCTTCCAGATCATGTGCTCCAAATGATGATTAAGCCATCAGTC 308
QY 1436 CACAGTGAAGCCGTTAATTAAGAGAGAAATGATGATTAATGAATCAAGTACAGAGAA 1495
DB 309 CACAGTGAAGCCGTTAATTAAGAGAGAAATGATGATTAATGAATCAAGTACAGAGAA 368
QY 1496 AGCAAGCAACTTTCAGAACAGCTAGAGTTTAAATCAACCTGCGCAAAATGTTTTGA 1555
DB 369 AGCAAGCAACTTTCAGAACAGCTAGAGTTTAAATCAACCTGCGCAAAATGTTTTGA 428
QY 1556 AAAACAGAACAAAGAGACAGAGCTGATGCTGCTGGCCAGAGCCAAAGGTTTCA 1615
DB 429 AAAACAGAACAAAGAGACAGAGCTGATGCTGCTGGCCAGAGCCAAAGGTTTCA 488
QY 1616 TGGCGAAATTAAGATTTGACAGATGCTGCTGACACAGGAGCGATCTTGGCATC 1675
DB 489 TGGCGAAATTAAGATTTGACAGATGCTGCTGACACAGGAGCGATCTTGGCATC 548
QY 1676 TAAACCGCTGAGAGTTTACCGGAAACAGCCAGAGAGAGCTTAAATGTCATATGGAAT 1735
DB 549 TAAACCGCTGAGAGTTTACCGGAAACAGCCAGAGAGAGCTTAAATGTCATATGGAAT 608
QY 1736 CTGTGCTGCTTTGAAGTAAAGAAACATATTAAGTCTGATGACAGAAAGCCAGCA 1795
DB 609 CTGTGCTGCTTTGAAGTAAAGAAACATATTAAGTCTGATGACAGAAAGCCAGCA 668

Qy 1796 GATGCTTG 1803
Db 669 GATGCTTG 676
RESULT 3
CK000929
LOCUS CK000929 815 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT_16363378 NIH_MGC_221 Homo sapiens cDNA clone
IMAGE 30708013 5', mRNA sequence.
ACCESSION CK000929
VERSION CK000929.1 GI:38526963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@db-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: NDAMI073 row: 1 column: 14
 High quality sequence stop: 613.
FEATURES
Source
 1..815
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30708013"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_221"
 /note="Organ: mixed; Vector: pYX-Anc; Site 1: EcoRI;
 Site 2: NotI; Library is oligo-dT primed and directionally
 cloned. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to RNA size fraction. Ligated with
 EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Anc vector. Average insert size
 4-5kb. Adaptor 5' (AATTCGACGACGAG) 3' and 5'd
 (CTGTGCGCCG) 3'. 3' linker sequence - GCGGCCGCTGAGAGC T18.
 Sequencing primers 3' end: T3 promoter primer 5'd
 (ATTACCTCTCACTTAAGGAG) 3'. 5' end: T7 promoter primer 5'd
 (TAATGACACTCACTAAGG) 3'. Library was constructed in the
 laboratory of M. Bento Soares. Note: this is a NIH_MGC
 library"
ORIGIN
 Query Match 30.4%; Score 618.4; DB 7; Length 815;
 Best Local Similarity 96.2%; Pied. No. 4.6e-139;
 Matches 685; Conservative 0; Mismatches 20; Indels 7; Gaps 5;
Qy 1122 TGACATCATCAGCAGACGCTTTGCTAAAGAAAGTAACAGNAGAGAGTGAACAACACT 1181
 |||||
Db 28 TGAATCATCAACAGCA- GCTTTGCTAAAGAAAGTAACAGNAGAGTGAACAACACT 86
 |||||
Qy 1182 GTTCTAGACCCCAAAAATGAGNCTGATATTGATAT- GCGTAGCGCTGAGNAGAGATCATC 1240
 |||||
Db 87 GTTCAAGACCCATTAATGAGAACTGAATTAATGAGATAGCCTGAGAGAGAGATCATC 146
 |||||

Qy 1241 AACAGACGATAAATGAGAGGAGTCTATTAGCCNTGGGTGAGTTCAACATGCGCTG 1300
 |||||
Db 147 AACAGACGATAAATGAGAGGAGTCTATTAGCCNTGGGTGAGTTCAACATGCGCTG 206
 |||||
Qy 1301 TATGAGCTCCTG--CATGCTGACACACACCGAGGGCTTGTCTAAGTGACAGAAACTGTT 1358
 |||||
Db 207 GATGAGCTCCTGAGTGGCTGACACACCGAGGGCTTGTCTAAGTGACAGAAACTGTT 266
 |||||
Qy 1359 GGAGAGACCTTAAGGAGTGAATTAAGTAACTTCCCAAGCATCTATGCTCCAAATGAT 1418
 |||||
Db 267 GGAGAGACCTTAAGGAGTGAATTAAGTAACTTCCCAAGCATCTATGCTCCAAATGAT 326
 |||||
Qy 1419 GTATTAGCCCATGAGTCCACAGTGAAGCCGTTATATAAGAGAGAAATGATCTATTGAA 1478
 |||||
Db 327 GTATTAGCCCATGAGTCCACAGTGAAGCCGTTATATAAGAGAGAAATGATCTATTGAA 386
 |||||
Qy 1479 TCAAGTCAAG 1538
 |||||
Db 387 TCAAGTCAAG 446
 |||||
Qy 1539 TGGCAAAATGTTTGGAAAAAACAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
 |||||
Db 447 TGGCAAAATGTTTGGAAAAAACAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 506
 |||||
Qy 1599 CAGGCCAAAGGCTTCATGGGAAATTGAGGATTGGAGAGTGGAGTGGAGTGGAGTGGAG 1658
 |||||
Db 507 CAGGCCAAAGGCTTCATGGGAAATTGAGGATTGGAGAGTGGAGTGGAGTGGAGTGGAG 566
 |||||
Qy 1659 CGTCAATCTGTTGGCATTAACCGCTGGAGAGTTTACCGGAAACAGCCAAAGAGAGAGTT 1718
 |||||
Db 567 CGTCAATCTGTTGGCATTAACCGCTGGAGAGTTTACCGGAAACAGCCAAAGAGAGAGTT 626
 |||||
Qy 1719 AATGTCATATGAGAAAGTCTGCTGCTTTGAGACTTAAGAAACAATTAAGAGTCTG 1778
 |||||
Db 627 AATGTCATATGAGAAAGTCTGCTGCTTTGAGACTTAAGAAACAATTAAGAGTCTG 686
 |||||
Qy 1779 ATGCAAGAAAGGCGAGAGATGCTTGAAGATGCCCAAAATGTGAGAGACA 1830
 |||||
Db 687 ATGCAAGAAAGGCGAGAGATGCTTGAAGATGCCCAAAATGTGAGAGACA 735
 |||||
RESULT 4
CR549848
LOCUS CR549848 764 bp mRNA linear EST 12-JUL-2004
DEFINITION DKFZp69K061_r1 469 (synonym: pk1d1) Pongo pygmaeus cDNA clone
IMAGE DKFZp69K061 5', mRNA sequence.
ACCESSION CR549848
VERSION CR549848.1 GI:50243472
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 1 (bases 1 to 764)
 Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German
 Cancer Research Center, Heidelberg/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp69K061) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/projects/cdna/.
FEATURES
 Location/Qualifiers
TITLE Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
JOURNAL This is the 5' sequence of the clone insert Clone from S. Wiemann,
COMMENT Molecular Genome Analysis, German Cancer Research Center (DKFZ);
 Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German
 Cancer Research Center, Heidelberg/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp69K061) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/projects/cdna/.
 Location/Qualifiers

```

source
1. .764
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469k061"
/lib_issue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkl1)"
/notes="Vector: pSport1_SfiI; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN

Query Match      27.7%; Score 561.6; DB 7; Length 764;
Best Local Similarity 96.4%; Pred. No. 2.9e-125;
Matches 585; Conservative 0; Mismatches 20; Indels 2; Gaps 1;

QY 1283 CAGTTCCAAATGCGCCCTGTATGAGCTCTG--CATGTGACACACACCGGCGCTTGCTA 1340
DB 1 CAGTTCCAAATGCGCCCTGTATGAGCTCTGACATGCTGACACACACCGGCGCTTGCTA 60
QY 1341 AGTAGAGAGAAACCTGTGTGAGAGAGACCTTAAAGCCATTGAATTGAACCTTGCAAGCAT 1400
DB 61 AGTAGAGAGAAACCTGTGTGAGAGAGACCTTAAAGCCATTGAATTGAACCTTGCAAGCAT 120
QY 1401 CATGTGCTCCAAATGATGATATTAGCCCATCAGTCCACAGTGGAAAGCCGTTAATAAGCA 1460
DB 121 CATGTGCTCCAAATGATGATATTAGCCCATCAGTCCACAGTGGAAAGCCGTTAATAAGCA 180
QY 1461 GGAATATATCTAATTGAAATCAAGTGCAGGAGAAAGAAAGCAACCTTCGAACAGCTA 1520
DB 181 GGAATATATCTAATTGAAATCAAGTGCAGGAGAAAGAAAGCAACCTTCGAACAGCTA 240
QY 1521 GAGGTTTAAATCAACGCTGCAAAATGTTTGGAAAAAAGAGAAAGAAAGAGAGAGAG 1580
DB 241 GAGGTTTAAATCAACGCTGCAAAATGTTTGGAAAAAAGAGAAAGAAAGAGAGAGAG 300
QY 1581 CTGATGATGCTCTTGGCCGAGCCAAAGGGTCTCATGCGCAAAATTGAGATTGCGAGAG 1640
DB 301 CTGATGATGCTCTTGGCCGAGCCAAAGGGTCTCATGCGCAAAATTGAGATTGCGAGAG 360
QY 1641 TGGCTGATGACACAGAGGCTCATCTTGTGGCATCTTAACCCGCTGGAGGTTTACCGGAA 1700
DB 361 TGGCTGATGACACAGAGGCTCATCTTGTGGCATCTTAACCCGCTGGAGGTTTACCGGAA 420
QY 1701 ACAAGCCAAAGAGAGCTTAAATGTCATATGAAAGTCTGCTGCTTTGAAGCTTAAGAA 1760
DB 421 ACAAGCCAAAGAGAGCTTAAATGTCATATGAAAGTCTGCTGCTTTGAAGCTTAAGAA 480
QY 1761 GAAACATATTAAGAGTCTGATGCAAGAAAGCCAGAGATGCTTGCAGAGATGCCCAAAATCT 1820
DB 481 GAAACATATTAAGAGTCTGATGCAAGAAAGCCAGAGATGCTTGCAGAGATGCCCAAAATCT 540
QY 1821 GCAGAGCAAAATTTGACCAAGACATTAATTAACCTTGAAGAAAAAAGAAAAAAGAAAAA 1880
DB 541 GCAGAGCAAAATTTGACCAAGACATTAATTAACCTTGAAGAAAAAAGAAAAAAGAAAAA 600
QY 1881 AAAANAC 1887
DB 601 ACCAAGC 607

RESULT 5
CK781134 752 bp mRNA linear EST 23-FEB-2004
LOCUS CK781134
DEFINITION U1-M-GVO-cln-d-16-0-U1.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30616311 5', mRNA sequence.
ACCESSION CK781134
VERSION CK781134.1 GI:42746812
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
1 (bases 1 to 752)
AUTHORS NIH-WGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.iowa.edu/distribution/mouseefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

FEATURES
source
Location/Qualifiers
1..752
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAG:30616311"
/lib_issue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GVO"
/notes="Organ: Brain; Vector: pYX-Anc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Anc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGCTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemm Chin, Ph.D., program coordinator."

ORIGIN

Query Match      25.9%; Score 526.8; DB 7; Length 752;
Best Local Similarity 82.2%; Pred. No. 8.2e-117;
Matches 617; Conservative 0; Mismatches 130; Indels 4; Gaps 1;

QY 99 TGGGTGATGATAGCTGAACAACAAGCTCATATAGATGAAGCAAAAGTGGCA-- 156
DB 1 TTGGAAGACTGATAGCTGAGCAGCAAGCCACATAGATGAAGCAAGACTGGCCAC 60
QY 157 --CAGTACTGATGAGCCCTGGGAAAGGCTTTCTTATCCAAAGAAAGTATGCGACGGC 214
DB 61 AGCTATTTGAACCTGAGCCCAAGGAGGATTTATTCAGAAAGATGTCGCGCGC 120
QY 215 ACACCTTTACAGTCACTTAAAGAAAGATGTCAGAAAGGCTGTCGCACTGATGAAG 274
DB 121 ACACCTTTACAGTCACTTAAAGAAAGATGTCAGAAAGGCTGTCGCACTGATGAAG 180
QY 275 CCATTTCTCAATCACTCACTGATGCAAGATGATGATGATGATGATGATGATGATGAT 334
DB 181 CCATTTCTCAATCACTCACTGATGCAAGATGATGATGATGATGATGATGATGATGAT 240
QY 335 GCATCTGGAAGAGTCTGAGGCGAGCCACCTCTATCTCTGCAAGCTTGAAGATGAAG 394
DB 241 GCATCTGGAAGAGTCTGAGGCGAGCCACCTCTATCTCTGCAAGCTTGAAGATGAAG 300
QY 395 AACAGATCAGTGAATTAAGAAATGTCAGTGAACATGGAAGAAAGCTACAGCGGTTATG 454
DB 301 AACAGATCAGTGAATTAAGAAATGTCAGTGAACATGGAAGAAAGCTACAGCGGTTATG 360
QY 455 AAACCTTTAAAGAGGAGGAGAAAGTATGCTGATGATGATGATGATGATGATGATGATG 514

```

Db 361 AGACTCTGAGGAGGAGGAGAGGATGATGCGCAGATCCGAGGAGCATGMAAAAGAGC 420

Qy 515 TATCTGCGCAAGCTCTTCAAGATAAGCTTGACCAATGCTTTTCATTGGAGAAATAC 574

Db 421 TATCTGCGAGAGCTGTATACAGATTAAGTTGACCAATGCTTTTCATTGGAGAAATAC 480

Qy 575 ACACACTGATGAGAGAGAGGAGGAGCAACTACTGATGATGATGAGAGCTAGCAGAAAGT 634

Db 481 ACACCTGTGTGAGAGAGGAGGAGGAGCAACTGCTGATGATGATGAGAGCTAGCAGAAAT 540

Qy 635 TCTGTGTGATCAGATGTCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 694

Db 541 TCTGTGTGATCAGATGTCATTTGATGATGATGATGATGATGATGATGATGATGATGAT 600

Qy 695 ACCTGGAAGATCTCTGGAATTTGATCTTCAAGTAAACCAACAGCAGAGAGCAGAGA 754

Db 601 ACCTGGAAGATCTCTGGAATTTGATCTTCAAGTAAACCAACAGCAGAGAGCAGAGA 660

Qy 755 CCATTAAGGAGAAATTAAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 814

Db 661 CAATCAGAGAGAGAGATGAG 720

Qy 815 CTGAAGTATGCGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845

Db 721 CAGAGCTCATGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 751

RESULT 6
BU701246 800 bp mRNA linear EST 15-JUL-2003
LOCUS BU701246
DEFINITION UI-M-EX0-bwv-a-04-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
IMAGE:5704587 5', mRNA sequence.
ACCESSION BU701246.1 GI:23624770
VERSION BU701246.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@db-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNM at: http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. 800
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5704587"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_id="NIH_BMAP_EX0"
/note="Organ: brain; Vector: pYX-asc; Site: 1; Site 2: Not 1; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction."

ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.9%; Score 486.4; DB 5; Length 800;
Best Local Similarity 81.3%; Pred. No. 5.4e-107;
Matches 610; Conservative 0; Mismatches 135; Indels 5; Gaps 4;

Qy 1095 TATCAACAGCAGATAG-AGTGTGACAGTCACTATCAACGACGCTTTTCTTAAAGAA 1153

Db 2 TATCAACAGCAGATAGAAATGAAAGGCTGAAACATCAAGCAGA-ACGTGTCTGAAAGAA 60

Qy 1154 AGTAAAGNAGAGTGAACAACACACTGTTCTAGACCAAAATGAGNCTGATTTGAT 1213

Db 61 AGTAAAGAGAGGAG 120

Qy 1214 AT-GGCTAGCTGAGAGAGAGAGATCATCAACAGAGATTAAGAGAGGCTCTAT 1272

Db 121 CTGGGAGAGAGCTGAGAGAGAGATTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

Qy 1273 AGCCTGAGTCACTTCAACATGCTGTATAGCTCT--GCATGCTGACACACACGA 1330

Db 181 AGCAGTGGGTGATTCAG 240

Qy 1331 GGGCTTGTAGTGAAG 1390

Db 241 AGGCTTGTAGTGAAG 300

Qy 1391 TGCCAGCATGATGCTGCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1450

Db 301 GGTAAAGATGATGCTGCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 360

Qy 1451 TAAATTAAG 1510

Db 361 CAACAAAG 420

Qy 1511 GAACAGCTAGAGTTTAAATCAACCTGCGCAAAATGTTTGAAGAGAGAGAGAGAG 1570

Db 421 GTAAAGCTAGAGATTTCAATCAACCTGCGCAAAATGTTTGAAGAGAGAGAGAGAG 480

Qy 1571 GAACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1630

Db 481 AAAGCAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

Qy 1631 TTTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1690

Db 541 TTTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

Qy 1691 TTTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750

Db 601 TCTCCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Qy 1751 AGTAAAG 1810

Db 661 CATCAAG 720

Qy 1811 CCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1840

Db 721 TCCCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 750

RESULT 7
BX470991 643 bp mRNA linear EST 04-SEP-2003
LOCUS BX470991
DEFINITION DK22686M17123 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
IMAGE:5704587 5', mRNA sequence.
ACCESSION BX470991

VERSION	EX470991.1	GI:31665308
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 643)	
AUTHORS	Bahr,A., Lauber,J., Mewes,H.W., weil,B., Amid,C., Osanger,A., Robo,G., Han,M. and Wiemann,S.	
TITLE	EST (Bahr,A., Lauber,J., Mewes,H.W., weil,B., et al.)	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: MIPS MIPS Inspektorat Landstr.1, D-85764 Neuherberg, Germany This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available. This clone (DKFP686M17123) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. Location/Qualifiers	
FEATURES	source	
	1..643	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="DKFP686M17123"	
	/dev_stage="adult"	
	/lab_host="DH10B"	
	/clone_lib="686 (synonym: hlc3)"	
	/note="Vector: pTriblex2; Site_1: SfIra; Site_2: SfIIB; cDNA-collection"	
ORIGIN		
Query Match	23.7%; Score 481; DB 5; Length 643;	
Best Local Similarity	91.0%; Pred.No.11e-105;	
Matches 588; Conservative	0; Mismatch 45; Indels 13; Gaps 7;	
Dp	722 CAGTAGTAAACACAGCAGCAGGAGAGACCATTAGGGAATAATAGATGCCTAC	781
OY	1 CAGTAGTAAACACAGCAGCAGGAGAGACCATTAGGGAATAATAGATGCCTAC	60
Dp	782 AGAAGAGCTGGATTAAGTTAATTAACCTAGGTTCGAACCTATGCCCATGTGGGAGC	841
OY	61 AGAAGAGCTGGATTAAGTTAATTAACCTAGGTTCGAACCTATGCCCATGTGGGAGC	120
Dp	842 CTGATTAACCCATGTCAAGAAGATATACATGATTAATTCAGCGATGGATCTCTA	901
OY	121 CTGATTAACCCATGTCAAGAAGATATACATGATTAATTCAGCGATGGATCTCTA	179
Dp	902 AATAAAGCTTGAAAAGACCCGATGACAACCTGAG-----AGCAATGCAGCTGCCGTC	955
OY	180 AATAAAGCTTGAAAAGACCCGATGACAACCTGAGCGCAATGCAGCTGCCGTTAG	239
Dp	956 AGTACAGATGACTGCAGGCGGTATTTTCTGGGTAGATATTCAGAGTGTAAAGTTCCGT	1015
OY	240 TACCAAGATGACTGCAGGCGGTATTTTCTGGGTAGATATTCAGAGTGTAAAGTTCCGT	299
Dp	1016 TCATATGTTCCATATTGGAACAGCATCTTGAAACCTGCAGCAGCAGATTTATGTCTAGA	1075
OY	300 TCATATGTTCCATATTGGAACAGCATCTTGAAACCTGCAGCAGCAGATTTATGTCTAGA	358
Dp	1076 GCAATTTAAGCTGAGGCTTATCAACAGCAGATG-AATGGATGAGCAGTCACTACAGC	1134
OY	359 GCAATTTAAGCTGAGGCTTATCAACAGCAGATGAAATGAAACAGTCACTACAGC	418
Dp	1135 AGAGCGTTTTCTAAAGAAAGTAAACAGNAGAGTGAACAAACACTTTCTAGACCAA	1194
OY	419 AGA-GCTTTTCTAAAGAAAGTAAACAGNAGAGTGAACAAACACTTTCTAGACCAA	477
Dp	1195 AAATGAGATGATATTGATAT-GGCTACCTGGAGNAGAGATCATCAACAGCAGATA	1253

Db	Query Match	Score	DB	Length	Matches	524;	Conservative	0;	Mismatches	59;	Indels	1;	Gaps	1;
Db	478	TAATGAACTGAAATTGATATGGATAGACCTCGGAGAGAGATATCAACAGACGACTA	537											
Qy	1254	AACGAGAGGGGCTCTATTAGCCNTGGGTCACTTCCAAATGCCCTGTATGAGCTCTG-	1312											
Db	538	AACGAGAGGGGCTCTATTAGCTTGGGTGAGTTCCAAATGCCCTGTATGAGCTCTG	597											
Qy	1313	-CATGCTGACACACACCGAGGGCTTGGTCAAGTGAAGCAAACTGT	1357											
Db	598	CATGCTGACACACACCGAGGGCTTGGTCAAGTGAAGCAAACTGT	643											
RESULT 8														
LOCUS	CO700994/c	589 bp	mRNA	linear	EST 26-JUL-2004									
DEFINITION	DG32-191b14 DG32-liver	Canis familiaris	cdna 3'	mRNA sequence.										
ACCESSION	CO700994													
VERSION	CO700994.1	GI:50649662												
KEYWORDS	EST													
SOURCE	Canis familiaris (dog)													
ORGANISM	Canis familiaris													
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.													
REFERENCE	1 (bases 1 to 589)													
AUTHORS	Schlueter, T., Hermanns, J., Weindel, M., Schuette, D., Kranz, H.,													
	Henrich, J., and Loebber, R.													
	Dog arrayTAG cDNA clone collection													
	Unpublished (2004)													
	Contact: Thomas Schlueter													
	LION bioscience AG													
	Walldorferstrasse 98, D-69123 Heidelberg, Germany													
	Tel.: +49 6221 4038 150													
	Fax: +49 6221 4038 290													
	Email: Thomas.Schlueter@lionbioscience.com.													
FEATURES														
SOURCE	Location/Qualifiers													
	1..589													
	/organism="Canis familiaris"													
	/mol_type="mRNA"													
	/strain="Beagle"													
	/db_xref="taxon:9615"													
	/tissue_type="liver"													
	/dev_stage="adult"													
	/lab_host="DH10B"													
	/clone_id="DG32-1liver"													
	/note="Organ: liver; Vector: Dog pBluescript LION"													
ORIGIN														
	Query Match	23.5%	Score	478.2;	DB	7;	Length	589;						
	Beet Local Similarity	89.74;	Pred. No.	5e-105;										
	Matches	524;	Conservative	0;	Mismatches	59;	Indels	1;	Gaps	1;				
Qy	176	GGGAAGGCTTTTCATCCAGAGAAAGATGATGGCAGCGACACCCCTTACAGTCAATTA	235											
Db	587	GGTAAGGCTTTTCATCCAGAGAAAGATGATGGCAGCGACACCC-TTACAGTCAATTA	529											
Qy	236	AAGAAGATGTCAAAAGCCGTGCTGTGCACTGATGAAGCCATTTCTCAATCAACTAGT	295											
Db	528	AAGAAGATGTCAAAAGCCGTGCTGTGCACTGATGAAGCCATTTCTCAATCAACTAGT	469											
Qy	296	TCCATGCAAGATGATCAAGTCTTTGAGAGCTGTGAACCGATGTGGAAAGCTGTAGGC	355											

Db 288 AGAATGATGCAAGATCTGAGGTAAGCAAGACATATCTGCCAAGCTGTTCAG 229

Qy 536 ATAGCTTGACCAATGGTTTTCATTTGGGAGAACATACACACTGTGTGAGAGAGG 595

Db 228 ATAGCTTGACCAATGGTTTTCATTTGGGAGAACATACACACTGTGTGAGAGAGG 169

Qy 596 AAGCAAACTCTGATGTGATGAGCTAGCAGAAAAGTTCTGGTGTGATCAGATGTCT 655

Db 168 AAGCAAACTCTGATGTGATGAGCTAGCAGAAAAGTTCTGGTGTGATCAGATGTCT 109

Qy 656 TGATAGTTACCTAAAGATATCTAAGATTTTCATCCGGACCTGGAAATCTGGAAATG 715

Db 108 TGTAGTAACTACTAAGATATACAGAGATTTTCATCCGGACCTGGAAATCTGGAAATG 49

Qy 716 ATCTCTCAGTGTGATGATGAGCTAGCAGAAAAGTTCTGGTGTGATCAGATGTCT 759

Db 48 ATCCCTCAGTGTGATGATGAGCTAGCAGAAAAGTTCTGGTGTGATCAGATGTCT 5

RESULT 9
CB164054/c 469 bp mRNA linear EST 30-JAN-2003

LOCUS K-EST0225079 L17N670205n1 Homo sapiens cDNA clone

DEFINITION L17N670205n1-37-C03 5', mRNA sequence.

ACCESSION CB164054

VERSION CB164054.1 GI:28150180

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 469)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

AUTHORS

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsun@gmail.kr
Plate: 37 row: C column: 03
High quality sequence stop: 469.

FEATURES
source
1. .469
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L17N670205n1-37-C03"
/sex="F"
/lab_host="Top10F"
/clone_lib="L17N670205n1"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN
Query Match 23.1%; Score 469; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 8.1e-103; Indels 0; Gaps 0;
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1369 CTAAGCCATTGAATTAAGCTTGGCAAGCATCTGCTCCAAAATATGATTAGCCC 1428

Db 469 CTAAGCCATTGAATTAAGCTTGGCAAGCATCTGCTCCAAAATATGATTAGCCC 410

Qy 1429 ATCAGTCACAGTGGAGCCGTTAATAAGCAGAAATGATCTAATTGAATCAAGTGAG 1488

Db 409 ATCAGTCACAGTGGAGCCGTTAATAAGCAGAAATGATCTAATTGAATCAAGTGAG 350

Qy 1489 GAGAAAGACAGAGACACCTTCAGAAACAAGCTAGAGGTTTAAATCAACGCTGGCAAAATG 1548

Db 349 GAGAAAGACAGAGACACCTTCAGAAACAAGCTAGAGGTTTAAATCAACGCTGGCAAAATG 290

Qy 1549 TTTTGAAAAAACAAGAAAGAAAGAGCAGCTGATGAGTGGCTTGGCCAGGCAAAAG 1608

Db 289 TTTTGAAAAAACAAGAAAGAAAGAGCAGCTGATGAGTGGCTTGGCCAGGCAAAAG 230

Qy 1609 GGTTCATGCGAAATTAAGAGATTGGCAGAGTGGCTGACTGACACGAGGCTCATCTGT 1668

Db 229 GGTTCATGCGAAATTAAGAGATTGGCAGAGTGGCTGACTGACACGAGGCTCATCTGT 170

Qy 1669 TGGCATCTAAACCCCTGGAGGTTTACCGGAACAGCCAGAGACGATTATGTCCATA 1728

Db 169 TGGCATCTAAACCCCTGGAGGTTTACCGGAACAGCCAGAGACGATTATGTCCATA 110

Qy 1729 TGGAGTCTGCTGCTGCTTGAAGCTAAAGAAACATATPAAGTCTGATGAGAAAG 1788

Db 109 TGGAGTCTGCTGCTGCTTGAAGCTAAAGAAACATATPAAGTCTGATGAGAAAG 50

Qy 1789 GCCAGCAGATGCTTGAAGATGCCCAAAATCTGCAGAGACAAATATTGA 1837

Db 49 GCCAGCAGATGCTTGAAGATGCCCAAAATCTGCAGAGACAAATATTGA 1

RESULT 10
BU139957 826 bp mRNA linear EST 25-NOV-2002

LOCUS BU139957

DEFINITION 603134131P1 CSEBCHL24 Gallus gallus cDNA clone CHEST115n8 5', mRNA sequence.

ACCESSION BU139957

VERSION BU139957.1 GI:25354439

KEYWORDS EST.

SOURCE Gallus gallus

ORGANISM Gallus gallus (chicken)

REFERENCE Boardman,P.E., Sans-Ezquerro,J., Overton,J.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE MEDLINE 1245392

JOURNAL PUBMED

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1. .826
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST115n8"
/dev_stage="16 day embryo"
/lab_host="DH10B"
/clone_lib="CSEBCHL24"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [stratagenel] vector to accommodate cDNA produced with the T-primed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BglI and

ORIGIN BamHI sites [5'ggccgcgtgcagcccgagatccgaaaaaag] [5'aattcttttccgagatccg93gctgcagc]

Query Match 23.0%; Score 468; DB 5; Length 826;

Best Local Similarity 79.1%; Pred. No. 1.6e-102; Matches 653; Conservative 0; Mismatches 164; Indels 9; Gaps 8;

QY 996 TTGACAGGTGGTAAATTGCTTCAATGCTTCAATTTGGAACAGACATCGAACTGTCAG 1055
 DB 2 TTGCTGGACATCAAGCTAGATCTATGTCCTCCCGTTGGAAC-GGATCTGGAGACAGTGAAG 60
 QY 1056 CACGACATTTGATGCTAGAGCAATTTAAGTCTGAGGCTTATCAACAGCAGATAG-AGTG 1114
 DB 61 CAACCAACTGAGGAACTTAAGCAATTTAAACAGAGCTTATCAGACAGCAGATAGAATG 120
 QY 1115 GTACGACTGACTATACCGAGACGCTTTGCTAAAGAAAGTAAACAGMAGAGTGCACAA 1174
 DB 121 GAAAGACTGAACATCAAGCAGA-ACATTTGCTGAAGAAAGTAAACAGGAAAGTGCACAA 179
 QY 1175 ACACACTGTTCTAGACCCAAAATGAGCTGATTTGATAT-GGCTAGCTGGAGNAGAG 1233
 DB 180 GCACACTGTTCAAGACCCCTCTCAGAGCTCAGAAATATGTGGACAGCCTGGAGGAAAA 239
 QY 1234 AATCATCAACAGACAGCATAACTGAGGGTGTCTATTAGCCNTGGCTCAGTTCCAA 1293
 DB 240 AATTATTAATAGACAGACAGCAAGCTGGAGGGTGTCTTTGATGCTTGGGAGCAGTTCCAGCA 239
 QY 1294 TGGCTCTGTATGACTCCT--GCATGCTGACACACACGAGGGCTTGT-AGTGGACAGA 1350
 DB 300 TGGCTCTGTATGACTCCT--GCATGCTGACACACACGAGGGCTTGT-AGTGGAGAACAGA 359
 QY 1351 AACCTGTGGAGAGACCTTAAGCCATTGAATTTGAATTTGCAAGCATCATGTGCTCC 1410
 DB 360 AACCTGTGGAGAGACCTTAAGCCATTGAATTTGAATTTGCAAGCATCATGTGCTCC 419
 QY 1411 AAAATGATGATTAGCCCATCACTGCAAGTGAAGCCGTTAAATAAGCAGAAATGATC 1470
 DB 420 AGATGATGTTTATGCTCATCAGCTCACTAGTAGAAGAGTTAAAGAAAGAGTGAAC 479
 QY 1471 TAATTGAATCAAGTGCAGAGAAAGAAAGCAACCTTCAAGCAACAGCTAGAGTTTAA 1530
 DB 480 TGAATGATCAAGTGCAGAGAAAGAAAGCAAGTAAATTTACGAGCAAGTGTGAACCTTCTGA 539
 QY 1531 ATCAAGCTGGGCAAAATGTTTGGAAAAAAGCAGAACAGAGACAGCTGATGCTG 1590
 DB 540 ATCAAGCTGGGCAAAATGTTTGGAAAAAAGCAGAACAGAGACAGCTGATGCTG 599
 QY 1591 CCTTGGGCGCAGGCAAAAGGTTCCATGCGCAAAATTGAGATTGCAAGCTGCTGACTG 1650
 DB 600 CCTTGGATCCAGGCGCAAGGTTTCCATGCTGAGTGAAGTATGCAAGCATGCTGACTG 659
 QY 1651 ACACGAGAGCTCATCTGTTGGCATCTTAACCGCTGGAGGTTTACCGGAAACAGCCAGG 1710
 DB 660 ACACGAGAGCTCATCTGTTGGCATCTTAACCGCTGGAGGTTTACCGGAAACAGCCAGG 719
 QY 1711 AGAGCTTAA-TGTCATATGAGAGCTGTGCTGCC-TTGAAGCTTAAAGAAACATTA 1768
 DB 720 AACAGCTTAAATTAATCCATATGAGAACTTGTGCTGCTTTTGAAGCCAAAGAAAGACATA 779
 QY 1769 TAAAGTCTGATGACAGAAAGCCAGCAGATGCTTGAAGATGCCCA 1814
 DB 780 TAGGTGTCTTAATGACAGACGCGCAGACAGATGCTTGAAGATGCCCA 825

RESULT 11
 CB525847 724 bp mRNA linear EST 09-JUN-2003
 LOCUS UI-M-FY0-cfe-p-07-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6848600 5', mRNA sequence.
 ACCESSION CB525847
 VERSION CB525847.1 GI:29359318
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 724)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaab-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.iowa.edu/distribution/mouse1.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

seq primer: pyx-5,
 Location/Qualifiers
 1..724
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6848600"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpd"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

FEATURES

source

QY 966 GACTGACGCGGTATTTTCTGGGTAGATATTCAGAGTGTGTAAGTTCGCTTCAATGCTTC 1025
 DB 1 GATTGACAGGAATATTCGATGCTGGTTCAGATGCGACGCAACAAATGAGCACATGCTTC 60
 QY 1026 CAATTGGAACAGCATTCGAAACCTGCAAGCAGACAGATGATGATGCTAGAGCAATTAG 1085
 DB 61 CCAATGGAACAG-ACCTGGAAGACGTCAGACAGCAGATTTGAAGCTGAAAGCAATTTAG 119
 QY 1086 TCTGAGCCTATCAACAGCAGATAG-AGTGTACGACTGACTCATCAGCAGCGCTTTT 1144
 DB 120 TCGAAGCTTATCAACAAGATGAAATGAAAGCTGACATCAACAGCA-ATTGTT 178
 QY 1145 GCTAAAGAAAGTACAGNAGAGTGCACAAACACTGTTCTTAGACCCAAAATGAGACT 1204
 DB 179 GCTGAAGAAAGTACAGAGAGGCGACAGACACAGTCCAGACCCGCTGTGAGAGCT 238
 QY 1205 GATATTGAAT-TGGCTAGCTGGAGNAGAGATCATCAACAGCAGCATTAATGAGAGG 1263
 DB 239 GAAGTTGATCTGGACAGCGCTGACAGCAATTTGACACAGACACACAGTTCAGAGG 298
 QY 1264 TGCTATTAGCCNTGGATCAATTCAGATGCTGATGAGCTCCT--GCATGCTGAC 1321

ORIGIN

Query Match 22.4%; Score 455.8; DB 6; Length 724;
 Best Local Similarity 81.0%; Pred. No. 1.4e-99;
 Matches 588; Conservative 0; Mismatches 132; Indels 6; Gaps 5;


```

Db      299 TGTCTCTTAGACAGTGGTTCAGTTCACAGACCCCTGGACGAGCTCTCGCATGGCTGAC 358
Qy      1322 ACACACCGAGGGCTTTGCTAAGTGACAGAAACCTGTTGGAGAGACCTTAAAGCCATTGA 1381
Db      359 ACACACCGAGGGCTTTGCTAAGTGACAGAAACCCCTTGGAGAGACCCGAAAGCCATTGA 418
Qy      1382 AATTGAACCTTSCAGACATCATGTGCTCAAAATGATATTTAGGCCATCATGTCAGAGT 1441
Db      419 GATTGAATGCTTAAGCATCACTGCTCCAAACGACGCTTTAGCCACAGTCCACGCT 478
Qy      1442 GGAAGCCGTTAATAAGCAGAAATGATCTAATTGAATCAAGTCAGAGAGAAAGCAAG 1501
Db      479 AGAAGCCGTCACAAAGCAGAAATGATCTCATGAGTCAGTGAAGCCGAAAGCAAG 538
Qy      1502 CAACCTTCAGAAACAGCTAGAGGTTTAAATCAACCTGCGAAATGTTTGGAAAAAAC 1561
Db      539 CAACCTTCAGTACAGCTCAGAAATTCATCAATCAGGCTGGAGGATATTTTGGAAAAAC 598
Qy      1562 AGAACAAGAGAGAGCAGCTGATGTCCTGGCCGACGAGGCAAAAGGTTCCATGGCGA 1621
Db      599 AGATCAAGAAAGAGCAGCTGTCGACAGCGCTTGGCCAGGCAAGGTTCCACGCTGA 658
Qy      1622 AATTGAGGATTTGACAGCAGTGGCTGACTGACACGAGCGTCATCTGTTGGCATCTAAAC 1681
Db      659 AATGAGATTTTGCAAGCTGGCTGACGAGACACAGAGGTCATCTCTTGGCATCTAAAC 718
Qy      1682 GCTGGG 1687
Db      719 TCTGGG 724

RESULT 12
LOCUS   CB723030              799 bp      mRNA      linear      EST 09-JUL-2003
DEFINITION
UI-M-GH0-cel-a-06-0-UI.r1 NIH BMAP_GH0 Mus musculus cDNA clone
IMAGE:6840943 5', mRNA sequence.
CB723030
CB723030.1 GI:29780172
EST.
Mus musculus (house mouse)
SOURCE  Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 799)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
        Email: cgabs-remail.nih.gov
        Tissue Procurement: Dr. Jim Lin, University of Iowa
        cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
        cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
        DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
        Clone Distribution: Distribution information can be found at
        http://genome.uiowa.edu/distribution/mouse1.html
        This clone was contributed by the Brain Molecular Anatomy Project
        (BMAP)
FEATURES
Source      1..799
             Location/Qualifiers
             1..799
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6"
             /db_xref="taxon:10090"
             /clone="IMAGE:6840943"
             /cissue="IMAGE:"whole brain"
             /dev_stage="1, 5, and 15 days newborn"
             /lab_host="DH10B (T1 phage resistant)"
             /clone_1lb="NIH BMAP GH0"
             /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
             Site 2: Not I; The library was constructed according
             to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
             1996. Denatured RNA was size fractionated on a 1% agarose

```

ORIGIN

gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGACTCAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): "gene discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

Query Match      22.1%; Score 449.8; DB 6; Length 799;
Best Local Similarity 76.5%; Pred. No. 4.2e-96;
Matches 615; Conservative 0; Mismatches 177; Indels 12; Gaps 5;

Qy      410 ATAAGATGTGTGCTAGTACATGGAAGGCTACAGCCGTTGTATGAACCTTTAAACGA 469
Db      1 ATAGAGCGTGTCCGTGACATGAGAAAGCTGCGCCGCTTACGACATCTTGAGCGAGA 60
Qy      470 GGGGAGAGAAATGATTGCTAGATCTGGGGGACTGATAAAGACATATCTGCCAAAGCTG 529
Db      61 GGGGAGAGAGATGATCGCCGATCCGAGGGCACTGAAAAAGAGTATCTGCCAGAGCTG 120
Qy      530 TTCAGATTAAGCTTGACCAAAATGCTTTTCATTTGGGAGAACATACACACTGCTGAAAG 589
Db      121 TACAGATTAAGTTGAGCAAAATGTTTCATCTGGGAAAGCATCACACGCTGTGTGAAG 180
Qy      590 AGAGGAGAGCCAAACTACTGATGTGATGAGAGTACGAGAAAGTTCTGTGTGATGACCA 649
Db      181 AACGGAAAGCCAACTGCTGATGTGATGAGAGTACGAGAAATTTCTGTGATGACCA 240
Qy      650 TGTCAATTGATGATTACCATTAAGATCTCAAGATTTTCATCCGGAGCTGGAAGATCTGTG 709
Db      241 TGTGCGTGTGTGTTACCATCAAGACACTCAAGATTTTCATCCGGAGCTGGAAGATCTGTG 300
Qy      710 GAATTGATCTTCACTGATGATAAACAACACGACGAGACGACGACATTAAGGGAAGAA 769
Db      301 GAATCGAACCCCTCGTGTGTAAGCAGACGAGAGTGCAGAGCAATCAAGAGAGGA 360
Qy      770 TAGATGACTACAGAGAGAGCTGATATAGTATTAACTAGTTCTGAATCTCAATTCGGG 829
Db      361 TTGACGCGCTGACAGAGAGAGCTGACATGTCATCAAGCTGCTCAGAGCTCATCGCTG 420
Qy      830 CATGTGGGAGCCCTGATTAACCCCATTTGTCAAGAGAGATATACATGATTAATTCAGCGA 889
Db      421 CATGTGGGAGCCCTGACAAACCATTTGTCAAGAGAGATTAAGAGACTTAATTCAGC-A 479
Qy      890 TGGGATTTCTTAATAAAGCTTGGAAAGACCGGATTAACCAACTTGA-----GAGCAAT 943
Db      480 TGGGATTTCTTAATAAAGCTTGGAAAGACCGGATTAAGAGAGAGAGGCGCATGCA 539
Qy      944 GCAGCTGCCGTCAGTACAGATGACCTGACGCGGATTTTCTGGGTAGATATTCAGAGT 1003
Db      540 GCTGCCGTTCAATCAACGAGAGAGTTCAGAGGAATATTCATCTGGGTTGATGACAGGC 599
Qy      1004 GGTAAAGTTGCTTCAATGCTTCCATTTGGAACAGACATCTCGAACTGTCAAGCAGAGAT 1063
Db      600 AACAAAGTTAGCAAAATGCTCNCATTTGAACAG-ACTGAGAGCCGTCAAGCAGAGAT 658
Qy      1064 TGATGTCTAGAGCAATTAATTAAGTCTGAGGCGCTTCAACAGCAGATAG-AGTGTAGCACT 1122
Db      659 TGAAGACTGAGCAATTAATTAAGTCTGAGAGCTTCAACAGATTAAGTAATTAAGTAAGGCT 718
Qy      1123 GACTCATACGACAGAGCCTTTTGTCTAAAGAAAGTAACAGNAGAGTGAACAACACTG 1182
Db      719 GAACCATCAGCAAACTGTGCTGAGAAAGTACAGAGAGGCGAG--ACAGCACAG 775
Qy      1183 TTCTAGACCCAAAATGAGCTGA 1206
Db      776 TCCAGAACCGGCTGATGAGCTGA 799

```

RESULT 13
BU217687 836 bp mRNA linear EST 25-NOV-2002
DEFINITION 603759087F1 CSQCHN04 Gallus gallus cDNA clone CHEST674n22 5', mRNA
SEQUENCE
ACCESSION BU217687.1 GI:25399501
VERSION BU217687.1
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE Boardman, P. E., Sanz-Bzquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
TITLE JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 836
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST674n22"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSQCHN04"
/note="Organ: whole embryo; Vector: Bluescript II KS(+);
Site1: EcorI; Site2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcorI, size-selected, and cloned into the NotI and EcorI
compatible sites of a custom modified MCS of the
Bluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
1996) 791, except that a significantly longer
reannealing hybridization was used."

Query Match 21.7%; Score 441.6; DB 5; Length 836;
Best Local Similarity 76.7%; Pred. No. 4.2e-96;
Matches 540; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

ORIGIN
232 ATTAAGAAGATGTCAAAAAGCGTGTGGCACTGGATGAGCCATTCTCAATCACT 291
|||||
1 ATTAAAGAAAGATGTCAAAAAGAGACACTGGCATGGATGAAGCCCTTCTCAGTGAAC 60
|||||
292 CAGTTCATGACGAAGATGATGATCCTTGAAGAGCTGGAAGCATCGTGAAGCTTGG 351
|||||
61 CAGTTCATGACGAAGATGATGATCCTTGAAGAGCTTGAAGCATGATGTAAGCTTGG 120
|||||
352 AGGACACCAACCTCTATCTCTGACAGAGGTGAGAAGATCAAGAGACAGATCACTGAAAT 411
|||||
Db 121 AGGACACCAACCTCTATCTCTGACAGAGGTGAGAAGATTAAGAGCAAAATTAAGTGAAT 180
|||||
412 AAGATGTGTCACTAGACATGAAAAAGCTAACCCGTTGTATGAAACTTTAAACAGAG 471
|||||

Db 181 AAGATGTGTCAATGATCTGCAAAAACCTTCCAGCCAGCTATGACGCTTAAGCAGCCA 240
|||||
472 GGAGAGGAAGAATGATGCTGATCTGGGGGAGCATGATTAAGACATATCTGCAAGAGTGT 531
|||||
241 GGGGAAGAAATGATGCTGCTGCTGACAGAGCAGATTAAGATGTATCTTAAAGTTGTG 300
|||||
532 CAGATTAAGCTTGACCAAAATGTTTTCATTGGGAGACATACACACTGTGTGAAGAG 591
|||||
301 CAAGATTAACCTAGACCAAAATGTCCTTATTGGCAAGATCATCAGACCTTGAAGGAA 360
|||||
592 AGGGAAGCCAAACTACTGATGATGATGAGCTTACGAGAAAGTTCTGTGTGATCAATG 651
|||||
361 AGGAGGCCAAATGTTGATGATGATGAGCAACTAGCGAAAGTTCTGTGTGATCAATG 420
|||||
652 TCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
|||||
421 GCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
|||||
712 ATTGATCTTCACTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
|||||
481 GTTGACCCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
|||||
772 GATGATCTACAGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 831
|||||
541 GATGATCTACAG 600
|||||
832 TGTGGGAG 891
|||||
601 TGTGGAG 660
|||||
892 GGATTCCTTAATTAAG 935
|||||
661 GATGCTTAAACAAACATGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704
|||||

RESULT 14
BU368661 793 bp mRNA linear EST 28-NOV-2002
LOCUS BU368661
DEFINITION 603597561F1 CSQCHN73 Gallus gallus cDNA clone CHEST567115 5', mRNA
SEQUENCE
ACCESSION BU368661
VERSION BU368661.1 GI:25876662
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE Boardman, P. E., Sanz-Bzquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
TITLE JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 793
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST567115"
/sex="female"
/tissue_type="not cerebrum or cerebellum"

```

/dev stage="adult"
/lab host="DH10B"
/clone_lib="CSECHN73"
/note="Organ: Brain; Vector: Bluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dt) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
Bluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) : 791, except that a significantly longer
reannealing hybridization was used."

```

ORIGIN

```

Query Match      21.7%; Score 440; DB 5; Length 793;
Best Local Similarity 75.4%; Pred. No. 1e-95;
Matches 576; Conservative 0; Mismatches 180; Indels 8; Gaps 2;

```

```

QY 244 GTCAAAAAGCGTGTGGCACTGATGAGCCATTCTCAATCACTCAGTCCATGAC 303
DB 1 GTCAAAAAGAGCTCTGCACTGATGAGCCCTTTCTCAGTGAACCAATTCATGAC 60
QY 304 AAGATGATGATCTCTGAGAGCCCTGAGACGATCTGGAACGCTTGAGGACGACCC 363
DB 61 AAGATGATGATCTCTGAGAGCCCTGAGACGATCTGGAACGCTTGAGGACGACCC 120
QY 364 TCTATCTCTGAGAGCTTGAGAGATCAAGAACGATCAGTGAATAATGATGTC 423
DB 121 TCAATCTCAAGAGGTTGAGAGATTAAGCAAAATTAATGATGATGTC 180
QY 424 GTAGACATGAAAAAGCTACAGCCGTTGATGAACTCTTAACAGAGGAGAGAAATG 483
DB 181 ATGATCTGAAAAAGCTACAGCCGTTGATGAACTCTTAACAGAGGAGAGAAATG 240
QY 484 ATTCTGATGATGAGGAGCTGATTAAGACATATCTGCAAAAGCTTTCAGATTAAGTT 543
DB 241 ATCTGCTGCTCAGAGGAGAGATTAAGATGATGATCTGTAAGTTGTCAGATTAAGTT 300
QY 544 GACCAATGCTTTTCTATTTGGAGAACATACACACTGCTGGAAGAGAGAGAGCCAA 603
DB 301 GACCAATGCTTTTCTATTTGGAGAACATACACACTGCTGGAAGAGAGAGCCAA 360
QY 604 CTACTGATGATGAGAGCTAGCAAAAAGTTCTGCTGATCAATGCTATGATGTT 663
DB 361 TTGTTGATGATGAGAGCTAGCAAAAAGTTCTGCTGATCAATGCTATGATGTT 420
QY 664 ACCATTAAGATCTCAGATTTTCACTCGGAGACCTGGAAGATCTGGAATGATCTTCA 723
DB 421 ACTATTAAGATCTCAGATTTTCACTCGGAGACCTGGAAGATCTGGAATGATCTTCA 480
QY 724 GTAGTAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
DB 481 GTAGTAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 784 GAGAGAGCTGATATGATTAATTAACCTAGGCTCTGAACCTATGCGGAGAGAGAGCT 843
DB 541 GAGAGAGCTGATATGATTAATTAACCTAGGCTCTGAACCTATGCGGAGAGAGCT 600
QY 844 GATTAACCCATTTCAAGAGAGATATCATGATTAATTAAGAGATGAGATTTCTTA- 902
DB 601 GATTAACCCATTTCAAGAGAGATATCATGATTAATTAAGAGATGAGATTTCTTA- 660
QY 903 -ATTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG- 958
DB 661 CAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG- 720
QY 959 ---ACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG- 999

```

DB 721 TACCAAGATGACTGACGACCAATATTTGATGCTGGTAGACATTGC 764

```

RESULT 15
LOCUS BQ215351 849 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT 7558465 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046743
5' mRNA sequence.
ACCESSION BQ215351
VERSION BQ215351.1 GI:20396751
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 849)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gga@bld-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM13293 row: a column: 16
High quality sequence stop: 685.
Location/Qualifiers

```

FEATURES

```

source
1..849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6046743"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

```

ORIGIN

```

Query Match      21.5%; Score 437.6; DB 5; Length 849;
Best Local Similarity 96.7%; Pred. No. 4e-95;
Matches 446; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

1427 CCATCAGTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATTAATTGAATCAAGTGC 1486

1 CCATCAGTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATTAATTGAATCAAGTGC 60

1487 AGGAG 1546

61 AGGAG 120

1547 TGTTTTGAAG 1606

121 TGTTTTGAAG 180

1607 AGGTTCCATGAG 1666

181 AGGTTCCATGAG 240

1667 GTTGGCATCTTAACCGCTGAGAGAGTTTACCGGAAACGCAAGAGAGAGAGTAAATGCCA 1726

241 GTTGGCATCTTAACCGCTGAGAGAGTTTACCGGAAACGCAAGAGAGAGAGTAAATGCCA 300

1727 TATGAGAGTGTGCTCCTTTGAAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1786

301 TATGAGAGTGTGCTCCTTTGAAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

Wed Jan 12 15:42:54 2005

us-10-089-887-46.rst

Page 12

OY	1787	AGGCAGAGAGTGCCTTGGAAATGCCCAAAATCTGCAGAGCAAAATTTCAGC	1846
Db	361	AGGCAGAGAGTGCCTTGCAGATGCCAAAATCTGCAGAGCAAAATTTCAGC	420
OY	1847	AAATTACTGTAAGAAAAAATAAAAAATAAAAAATAAAAAATC	1887
Db	421	AAATTACTGTAAGAAAAAATGGAAATCGTGGAAATC	461

Search completed: January 12, 2005, 12:49:27
Job time : 5757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 05:49:47 ; Search time 925 Seconds
(without alignments)
11526.026 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ttgcgcacgcgcacgaaca.....tngagagcccccacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N GeneSeq 23Sep04:*

1: geneSeq1980s:*

2: geneSeq1990s:*

3: geneSeq2000s:*

4: geneSeq2001as:*

5: geneSeq2001bs:*

6: geneSeq2002as:*

7: geneSeq2002bs:*

8: geneSeq2003as:*

9: geneSeq2003bs:*

10: geneSeq2003cs:*

11: geneSeq2003ds:*

12: geneSeq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1983	97.6	2028	6	ABK27463 DNA encod
2	1656.8	81.6	9274	6	ABK13324 DNA encod
3	1653.6	81.4	16684	12	AD024870 DNA encod
4	1637.6	80.6	3307	11	ADM02768 Human cal
5	1010	49.7	6006	4	AAH41086 Human cal
6	655.4	32.3	6007	8	ACA56949 Human adi
7	655.4	32.3	16122	4	AAS22908 DNA encod
8	655.4	32.3	16341	4	AAS23002 DNA encod
9	501	24.7	2181	6	ABK34676 Human CDN
10	497.8	24.5	2113	4	AAS29811 Human cyt
11	497.4	24.5	5453	4	AAS29791 Human cyt
12	438.8	20.8	547	4	AAI22375 Human bre
13	422.8	20.8	531	4	AAI13506 Human bre
14	357	17.6	401	5	AAE65919 Novel hum
15	352.2	17.3	4327	4	AAK51484 Human pol
16	352.2	17.3	4438	4	AAK51483 Human pol
17	266.8	13.1	4495	4	AAI58467 Human pol
18	266.8	13.1	4495	5	ADQ98882 DNA encod
19	266.8	13.1	4495	9	ADB48442 Novel hum
20	266.8	13.1	4578	4	AAK52468 Human pol
21	266.8	13.1	4578	4	AAK52467 Human pol

22	266.8	13.1	4775	4	AAI60253 Human pol
23	262.8	12.9	344	6	ABN75109 Human str
24	250.2	12.3	278	3	AAC32058 Human sec
25	235.6	11.6	1721	6	ABBS1546 Human CDN
26	216.2	10.6	4721	5	AAS84613 DNA encod
27	177.8	8.8	3684	10	ADA52704 Human cod
28	162.2	8.0	12995	4	ABL20575 Drosophi1
29	162.2	8.0	17678	4	ABL19341 Drosophi1
30	162.2	8.0	18011	4	ABL20689 Drosophi1
31	162.2	8.0	27845	4	ABL20591 Drosophi1
32	153.6	7.6	388	2	AAV90172 EST clone
33	135	6.6	39969	4	AAK85853 Human imm
34	134.6	6.6	563	12	ACH71003 Human gen
35	133.6	6.6	171	12	ACH84703 Human gen
36	115.8	5.7	31741	5	ABA21231 Human ner
37	93.8	4.6	1024	3	AAZ97148 Human pro
38	88.6	4.4	449	4	AAS29810 Human cyt
39	88.6	4.4	4156	4	AAS29790 Human cyt
40	87.4	4.3	22137	4	ABL20574 Drosophi1
41	87.4	4.3	48718	4	ABL20590 Drosophi1
42	87.4	4.3	63284	4	ABL20688 Drosophi1
43	87.4	4.3	71962	4	ABL19340 Drosophi1
44	82.6	4.1	3372	4	ABL1494 Human inh
45	69.4	3.4	1613	3	AAA07277 Human PST

ALIGNMENTS

RESULT 1
ID ABK27463 standard; CDNA, 2028 BP.
XX
AC ABK27463;
XX
DT 09-APR-2002 (first entry)
XX
DE DNA encoding colon cancer associated protein, CATX-14.
XX
KW Human; colon cancer; tumour; abnormal cell growth; melanoma;
KW cervical cancer; colorectal adenocarcinoma; Wilms' tumour; leukaemia;
KW lymphoma; antisense therapy; CATX; gene; ss.
XX
OS Homo sapiens.
XX
PN WO20011047-A2.
XX
PD 15-FEB-2001.
XX
PF 08-AUG-2000; 2000WO-US021606.
XX
PR 09-AUG-1999; 99US-0147933P.
XX
PA (FARB) BAYER CORP.
XX
PI Bowman BM, Wang K;
XX
DR WPI, 2002-121548/16.
XX
P-PSDB; AAU81310.
XX
PT New isolated nucleic acid involved in growth regulation in human colonic
PT epithelial cells, termed CATX, for diagnosing and treating abnormal cell
PT growth, and for use as a probe/primer for detecting tumors.
XX
PS Claim 13; Page 122; 130pp; English.
XX
CC The invention relates to an isolated nucleic acid (I) involved in growth
CC regulation in human colonic epithelial cells, termed CATX. (I) is useful
CC as a probe/primer for detecting tumours, preferably colon cancer. The
CC nucleic acid, encoded polypeptide and antibody are useful in diagnosis
CC and treatment of abnormal cell growth (such as cervical cancer,
CC melanomas, colorectal adenocarcinomas, Wilms' tumour, leukaemia and
CC lymphomas), in screening assays for the treatment of abnormal cell

Db 1857 AAAAAAAAAAAAAAAAAAAAAAAAAANACGCCGCTGATTTAGATTCAGCGCCGCT 1916
 Oy 1920 GAATTTGAGACCTGCGCGCGCGCGCTCGAGCCCTTAGTGAAGCTTATTTAGATGGA 1979
 Db 1917 GAATTTGAGACCTGCGCGCGCGCGCTCGAGCCCTTAGTGAAGCTTATTTAGATGGA 1976
 Oy 1980 ATCACTAGTCCGCGCGCTGAGAGTGCATATNGAGAGCTCCCAAGCGCT 2031
 Db 1977 ATCACTAGTCCGCGCGCGCTGAGAGTGCATATNGAGAGCTCCCAAGCGCT 2028

RESULT 2
 ABK13224
 ID ABK13224 standard; cDNA; 9274 BP.
 XX
 AC ABK13224;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE DNA encoding novel signal pathway protein protein PN9109.
 XX
 KM MAPKAP-K3, Ap-3 delta; APP-695; amyloid A4; Hep8; heat shock protein 8;
 KM L130; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDR9;
 KM cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;
 KM amphiphysin; PN9109; KIAA1106; neurodegenerative disorder;
 KM non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;
 KM inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;
 KM atherosclerosis; cardiac hypertrophy; hypoxic brain injury;
 KM yeast two-hybrid; signal transduction pathway; human; ss;
 KM mitogen activated protein kinase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..8508
 FT /*tag= a
 FT /product= "PN9109"
 FT /partial
 FT /note= "No start codon given"
 XX
 PN W0200198524-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 21-JUN-2001; 2001WO-US019762.
 XX
 PR 22-JUN-2000; 2000US-0213245P.
 XX
 PA (MYRI-) MYRIAD GENETICS INC.
 XX
 PI Heichman K, Bartel PL;
 XX
 DR WP1; 2002-122287/16.
 DR P-PSDB; AAU75100.
 XX
 PT New protein complexes comprising protein-protein interactions (e.g.
 PT MAPKAP-K3/AP-3 delta or C-NAP-1/Clathrin HC), useful for diagnosing
 PT physiological generative disorders or screening drugs for these diseases.
 XX
 PS Example 12; Page 27-31; 60pp; English.
 XX
 CC The invention describes an isolated protein complex, comprising two
 CC proteins. The protein complex comprises: protein kinase MAPKAP-K3 and Ap-
 CC 3 adaptor complex delta sub-unit; MAPKAP-K3 and amyloid A4 precursor
 CC protein, APP-695; MAPKAP-K3 and heat shock protein (Hsp) 8; leucine rich
 CC L130 and NY-REN-58; P38 Alpha and P38 Beta; protein kinase ERK3 and
 CC KIAA0934 (unknown function); ERK3 and cell cycle dependent kinase (CDK) 9
 CC ; ERK3 and protein kinase CLK; C-NAP-1 and Clathrin heavy chain; C-NAP-1
 CC and Amphiphysin; C-NAP-1 and novel protein PN9109 or C-NAP-1 and KIAA1106
 CC (unknown function) interactions. The protein complexes are useful for
 CC diagnosing physiological generative disorders, drug screening for agents
 CC that modulate the interaction of the proteins (thus identify drug
 CC targets), and identifying additional proteins in the pathway common to

CC the proteins. These physiological disorders include non-insulin dependent
 CC diabetes mellitus (NIDDM), neurodegenerative disorders (e.g. Alzheimer's
 CC disease), inflammatory diseases (e.g. rheumatoid arthritis and
 CC inflammatory bowel disorder) and other human disease such as
 CC atherosclerosis, cardiac hypertrophy and hypoxic brain injury. This
 CC sequence encodes the novel protein PN9109, found to bind C-NAP1 in a
 CC yeast two-hybrid assay used to determine components of signal
 CC transduction pathways, described in the method of the invention
 XX

Sequence 9274 BP; 3055 A; 1878 C; 2241 G; 2100 T; 0 U; 0 Other;
 SQ

Query Match 81.6%; Score 1656.8; DB 6; Length 9274;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1805; Conservative 0; Mismatches 72; Indels 18; Gaps 9;

Oy 8 GCTGACAGAAACACATCATCTCTCAGC-TCGCCGCCAGCCCTTGAATATGATAC 66
 Db 3612 GCTGACAGAAACACATCATCTCTCAGCCTCCGCCAGCCCTTGAATATGAAAC 3671
 Oy 67 TCTAAGCAGAGAGAGAAACATCGGCACTGCGTGAATGATGATGTAACACAGACC 126
 Db 3672 TCTAAGCAGAGAGAGAAACATCGGCACTGCGTGAATGATGATGTAACACAGACC 3721
 Oy 127 TCATATAGATTAAGTGAACAAACTGG-CCACAGTACTGATGAGCCCTGGGGAAG 182
 Db 3732 TCATATAGATTAAGTGAACAAACTGGGCCACAGTTACTGAAATTTGAGCCCTGGGGAAG 3791
 Oy 183 CTTTCTATCCAGAGAGATGATGTGGACGCCGACCCCTTTACAGTACATTAAGAAGA 242
 Db 3792 CTTTCTATCCAGAGAGATGATGTGGACGCCGACCCCTTTACAGTCAATTAAGAAGA 3851
 Oy 243 TGTCAAAAGGCTCTGTGGACCTGAGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 302
 Db 3852 TGTCAAAAGGCTCTGTGGACCTGAGATGAAGCCATTTCTCAATCACTCAGTTCCATGA 3911
 Oy 303 CAAGATAGATCAGATCTCTGAGACCTGGAAACGATGATGAGAGCTGAGAGCCAGCC 362
 Db 3912 CAAGATAGATCAGATCTCTGAGACCTGGAAACGATGATGAGAGCTGAGAGCCAGCC 3971
 Oy 363 CTCTATCTCTGAGAGGTTGAAAGATCAAGAAACAGATCAGTAAATAAATGATGTC 422
 Db 3972 CTCTATCTCTGAGAGGTTGAAAGATCAAGAAACAGATCAGTAAATAAATGATGTC 4031
 Oy 423 AGTAGACATGGAAGGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 482
 Db 4032 AGTAGACATGGAAGGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGAGAGAAAT 4091
 Oy 483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 542
 Db 4092 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCT 4151
 Oy 543 TGACCAATGGTTTTCATTGGGAGAAATATACACACTGGTGGAGAGAGAGCCAA 602
 Db 4152 TGACCAATGGTTTTCATTGGGAGAAATATACACACTGGTGGAGAGAGAGCCAA 4211
 Oy 603 ACTACTGATGTATGAGAGCTAGACAGAAAGTTCTGGTGTGATCACTGTCATTGATGAT 662
 Db 4212 ACTACTGATGTATGAGAGCTAGACAGAAAGTTCTGGTGTGATCACTGTCATTGATGAT 4271
 Oy 663 TACCATTAAGATATCTCAAGATTTTCATCCGGGAGCTGGAAGATCTGGAATTTGATCTTC 722
 Db 4272 TACCATTAAGATATCTCAAGATTTTCATCCGGGAGCTGGAAGATCTGGAATTTGATCTTC 4331
 Oy 723 AGTAGTAACCAACAGACAGAGAGAGACCATTAAGGAGAGAAATAGATGAGACTACA 782
 Db 4332 AGTAGTAACCAACAGACAGAGAGAGACCATTAAGGAGAGAAATAGATGAGACTACA 4391
 Oy 783 GGAGAGCTGATATAGTTATTAACCTAGAGTTCTGAACCTCATTCGGGATGCGGAGCC 842
 Db 4392 GGAGAGCTGATATAGTTATTAACCTAGAGTTCTGAACCTCATTCGGGATGCGGAGCC 4451
 Oy 843 TGATTAACCAATGTCAGAGAGATATATGATTAATTCAGCGATGGGATCTCTTA 902

Db 4452 TGATTAACCATTTGTCAGAGAGATAGATGATTAATTAACG-ATGGGATTTCTCTAA 4510
QY ATAAAGCTTGAGAAAGCCGATTTGACCACTTGAG-----AGCAATGACAGTGCCTGCA 956
Db 4511 ATTAAGTTTGAAGAGCCGATTTGACCAACTTGAAGAGGCAATGACAGGCTGCGCTTCACT 4570
QY 957 GTACAGATGAGCTGACAGCCGTAATTTTCTGGGTAGATATTTGAGGTGTGTTGCTT 1016
Db 4571 ACCAGATGAGCTGACAGCCGTAATTTGAGTGGGTAGATATTTGAGGTGTGTTGCTT 4630
QY 1017 CAATGCTCCAAATTTGAGAGCAAGCACTGCAAGTGTGAGAGAGATTTGATGCTTGAAG 1076
Db 4631 CAATGCTCCAAATTTGAGAGCAAGCACTGCAAGTGTGAGAGAGATTTGATGCTTGAAG 4689
QY 1077 CAATTTAGTCTGAGGCTTATCAACAGCAGATAG-AGTGTAGCAGTCACTCAACGCA 1135
Db 4690 CAATTTAGTCTGAGGCTTATCAACAGCAGATAGAGATTAAGATGATTCATCAAGCA 4749
QY 1136 GACGCTTTTGTCTAAAGAAATTAACAGNAGAGTGAACAAACACACTGTTCTAGACCCAA 1195
Db 4750 GA-GCTTTTGTCTAAAGAAATTAACAGNAGAGTGAACAAACACACTGTTCTAGACCCATT 4808
QY 1196 AATGAGCTGATTTGATAT-GGCTAGCCCTGAGAGAGAGATGATCAACAGACAGCATTA 1254
Db 4809 AATGAGCTGATTTGATATGAGATGAGCTGAGAGAGAGATGATCAACAGACAGCATTA 4868
QY 1255 ACTGAGGGTCTCTATTAGCCCTGAGTCAATGCTGATGAGCTCTG-- 1312
Db 4869 ACTGAGGGTCTCTATTAGCCCTGAGTCAATGCTGATGAGCTCTG-- 4928
QY 1313 CATGCTGACACACAGAGGGCTTGTCTAAGTACAGAAACCTGTTGAGAGAGACCTTAA 1372
Db 4929 ATGGCTGACACACAGAGGGCTTGTCTAAGTACAGAAACCTGTTGAGAGAGACCTTAA 4988
QY 1373 AGCCATTGAAATGAACTTGCCCAAGCATCATGCTCCAAATAGATATTAACCCATCA 1432
Db 4989 AGCCATTGAAATGAACTTGCCCAAGCATCATGCTCCAAATAGATATTAACCCATCA 5048
QY 1433 GTCCACAGTGAAGCCGTTAATTAAGCAGAGAAATGATCTAATTTGAATCAAGTCAAGAGA 1492
Db 5049 GTCCACAGTGAAGCCGTTAATTAAGCAGAGAAATGATCTAATTTGAATCAAGTCAAGAGA 5108
QY 1493 AGAAGCAAGCACTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGCGCAAAATGTTTT 1552
Db 5109 AGAAGCAAGCACTTCAAGCAAGCTAGAGGTTTAAATCAACGCTGCGCAAAATGTTTT 5168
QY 1553 GGAAGAAACAGAACAAAG 1612
Db 5169 GGAAGAAACAGAACAAAG 5228
QY 1613 CCATGCGCAAAATGAGATTTGACAGAGTGTGAGTGAACAGAGAGAGTCACTGTTGGC 1672
Db 5229 CCATGCGCAAAATGAGATTTGACAGAGTGTGAGTGAACAGAGAGAGTCACTGTTGGC 5288
QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1732
Db 5289 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 5348
QY 1733 AGCTGTGTGCTCTTGAAGCTTAAGAGAAATTAAGAGTCTGATGAGAGAGAGAGAGAGAG 1792
Db 5349 AGCTGTGTGCTCTTGAAGCTTAAGAGAAATTAAGAGTCTGATGAGAGAGAGAGAGAGAG 5408
QY 1793 GCGAGATGCTTGAAGATGCGCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1852
Db 5409 GCGAGATGCTTGAAGATGCGCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5468
QY 1853 CTTGAAAAAAGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1887
Db 5469 CTTGAAAAAAGGCTGCTGAG 5503

RESULT 3
ADQ24870

ID ADQ24870 standard; DNA; 16684 BP.
XX
AC ADQ24870;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7690.
DE
XX soft tissue sarcoma; cytosolic; gene therapy; vaccine; screening; human;
KM de.
XX
OS Homo sapiens.
XX
PN MO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003MO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7690; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytosolic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 16684 BP; 5630 A; 3109 C; 3754 G; 3814 T; 0 U; 377 Other;
XX
Query Match 81.4%; Score 1653.6; DB 12; Length 16684;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 74; Indels 18; Gaps 9;
QY 8 GGTGACAGAAACATCAATCATCTCTCAGC-TCGCCGCCAGCCCTTGAATATGATAC 66
Db 10732 GGTGACAGAAACATCAATCATCTCTCAGCTTCCGCCAGCCCTTGAATATGAAAC 10791
QY 67 TCTAAGGCGAGCGAGAGAAACATGCGCACTGCTGATGATGCTGAAACACAAGCC 126
Db 10792 TCTAAGGCGAGCGAGAGAAACATGCGCACTGCTGATGATGCTGAAACACAAGCC 10851
QY 127 TCATATAGATTAAGTGAACAAACTGG-----CCACAGTACTGAATAGCCCTGGGGAGG 182
Db 10852 TCATATAGATTAAGTGAACAAACTGGGGACACAGTTACTGGAATTAGCCCTGGGGAGG 10911
QY 183 CTTTCTATCAAGAGAGATATGTCAGCCGACACCCCTTACAGTCACTTAAGAGAG 242
Db 10912 CTTTCTATCAAGAGAGATATGTCAGCCGACACCCCTTACAGTCACTTAAGAGAG 10971
QY 243 TGTCAAAAAGCGTCTGCTGCGCACTGAGTGAAGCCATTCTCAATCACTCACTGTTCCATGA 302
Db 10972 TGTCAAAAAGCGTCTGCTGCGCACTGAGTGAAGCCATTCTCAATCACTCACTGTTCCATGA 11031

QY 303 CAAGATGATCAGATCCTTGAGAGCTGGAAACGATCGTGGAAAGCTCTGAGAGCCACACC 362
 Db 11032 CAAGATGATCAGATCCTTGAGAGCTGGAAACGATCGTGGAAAGCTCTGAGAGCCACACC 11091
 QY 363 CTCTATCTCTGACAGGTTGAGAAATCAAGAAACAGATCAGTAAAAATAGATGTGTC 422
 Db 11092 CTCTATCTCTGACAGGTTGAGAAATCAAGAAACAGATCAGTAAAAATAGATGTGTC 11151
 QY 423 AGTAGACATGAAAAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGGGAGAGAAAT 482
 Db 11152 AGTAGACATGAAAAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGGGAGAGAAAT 11211
 QY 483 GATTGCTAGATCTGGGGGAGTGTATTAAGATATCTGCCAAACCTGTTCCAGTAAAGT 542
 Db 11212 GATTGCTAGATCTGGGGGAGTGTATTAAGATATCTGCCAAACCTGTTCCAGTAAAGT 11271
 QY 543 TGACCAATGGTTTCAATTTGGAGAAACATACACACTGGTGGAAAGAGAGGACCA 602
 Db 11272 TGACCAATGGTTTCAATTTGGAGAAACATACACACTGGTGGAAAGAGAGGACCA 11331
 QY 603 ACTACGTAGTGTAGTGAAGCTAGCAGAAAAGTTCTGTGTATCAATGTCAATTTAGT 662
 Db 11332 ACTACGTAGTGTAGTGAAGCTAGCAGAAAAGTTCTGTGTATCAATGTCAATTTAGT 11391
 QY 663 TACCATTTAAAGATCTCAAGTTTCAATCCGGAGCTGGAAAGATCTCTGAAATTTCTTC 722
 Db 11392 TACCATTTAAAGATCTCAAGTTTCAATCCGGAGCTGGAAAGATCTCTGAAATTTCTTC 11451
 QY 723 AGTAGTAAAAACAAGCAGAAAGCAGCAGAGACCATTAAGGAGAAATAGATGAGACTCA 782
 Db 11452 AGTAGTAAAAACAAGCAGAAAGCAGCAGAGACCATTAAGGAGAAATAGATGAGACTCA 11511
 QY 783 GGAGAGCTGAGATAGTATTATTAACCTAGGTTCTGAATCTCATTCGGCATGTGGAGACC 842
 Db 11512 GGAGAGCTGAGATAGTATTATTAACCTAGGTTCTGAATCTCATTCGGCATGTGGAGACC 11571
 QY 843 TGATTAACCCATTTCTCAAGAAAGATATACATGATTAATTCAGCAGATGGCATTTCTCA 902
 Db 11572 TGATTAACCCATTTCTCAAGAAAGATATACATGATTAATTCAGCAGATGGCATTTCTCA 11630
 QY 903 ATAAAGCTTGAAGAACCGAGATTGACCAACTTGAG-----AGCAATGCAGCTGCCGTC 956
 Db 11631 ATAAAGCTTGAAGAACCGAGATTGACCAACTTGAG-----AGCAATGCAGCTGCCGTC 11690
 QY 957 GTACAGATGAGCTGACAGCCGTTATTTTCTGGTAGATATTGCAAGTGTGTAAGTTCGCTT 1016
 Db 11691 ACAGAGATGAGCTGACAGCCGTTATTTGACTGGTAGATATTGCAAGTGTGTAAGTTCGCTT 11750
 QY 1017 CAATGTCTCCAAATTGGAACAGCATCTGAAACTGTCAAGCAGCATGATGTGCTAGAG 1076
 Db 11751 CAATGTCTCCAAATTGGAACAG-ATCTGAAACTGTCAAGCAGCATGATGGAAGACTTAAG 11809
 QY 1077 CAATTTAAGTGTGAGGCTATCAACAGAGTAG-AGTGTAGCACTGAGCTCATCAAGCA 1135
 Db 11809 CAATTTAAGTGTGAGGCTATCAACAGAGTAG-AGTGTAGCACTGAGCTCATCAAGCA 11869
 QY 1136 GACGCTTTTGTAAAGAAAGTACAGNAGAGTGAACAACAACCTGTTCTAGACCCCAA 1195
 Db 11870 GACGCTTTTGTAAAGAAAGTACAGNAGAGTGAACAACAACCTGTTCTAGACCCCAA 11928
 QY 1196 AATGAGNCTGATATTGATAT-GGCTAGCTTGAAGNAGATCATCAACAGACAGCATTA 1254
 Db 11929 AATGAGNCTGATATTGATATGAGATAGCTTGAAGNAGATCATCAACAGACAGCATTA 11988
 QY 1255 ACTGAGAGGAGTCTCATATTAGCCNTGGGTCAGTTCCAAATGCGCTGTAGAGCTCCG-- 1312
 Db 11989 ACTGAGAGGAGTCTCATATTAGCCNTGGGTCAGTTCCAAATGCGCTGTAGAGCTCCG-- 12048
 QY 1313 CATGCTGACACACACGAGGGCTTGTAGTGAAGCAAAAATCTGTTGAGAGAGACCTTA 1372
 Db 12049 ATGCTGACACACACGAGGGCTTGTAGTGAAGCAAAAATCTGTTGAGAGAGACCTTA 12108

QY 1373 AGCATTAAGAAATTGAACCTTGCAAGACATATGCTCCAAAATATGATGATTAAGCCATCA 1432
 Db 12109 AGCATTAAGAAATTGAACCTTGCAAGACATATGCTCCAAAATATGATGATTAAGCCATCA 12168
 QY 1433 GTCCACAGTGAAGCCGTTATTAAGCAGGAAATGATCTAATTTGAATCAAGTCAGAGAGA 1492
 Db 12169 GTCCACAGTGAAGCCGTTATTAAGCAGGAAATGATCTAATTTGAATCAAGTCAGAGAGA 12228
 QY 1493 AGAAGCAAGCACTTCAAGAACAGCTAGAGCTTTAAATCAACGCTGGCAAAATGTTTT 1552
 Db 12229 AGAAGCAAGCACTTCAAGAACAGCTAGAGCTTTAAATCAACGCTGGCAAAATGTTTT 12288
 QY 1553 GGAAGAAAAACAAGCAAGAGAGCAGCTGATGTGTGCTTGGCCAGGCAAGGATT 1612
 Db 12289 GGAAGAAAAACAAGCAAGAGAGAGCAGCTGATGTGTGCTTGGCCAGGCAAGGATT 12348
 QY 1613 CCATGGCGAAATTGAGATTTTGACAGAGTGGCTGACTGACACGAGAGCTCATCTGTTGC 1672
 Db 12349 CCATGGCGAAATTGAGATTTTGACAGAGTGGCTGACTGACACGAGAGCTCATCTGTTGC 12408
 QY 1673 ATCTAAACCGCTGGAGAGTTTACCGGAAACAGCCAGAGAGAGCTTAATGTCATATGA 1732
 Db 12409 ATCTAAACCGCTGGAGAGTTTACCGGAAACAGCCAGAGAGAGCTTAATGTCATATGA 12468
 QY 1733 AGTGTGCTGCTTGAAGCTAAGAAAGAAACATATTAAGTGTGATGACAGAAAGGCA 1792
 Db 12469 AGTGTGCTGCTTGAAGCTAAGAAAGAAACATATTAAGTGTGATGACAGAAAGGCA 12528
 QY 1793 GCAGATGCTTGAAGATGCCCCAAATCTGACAGAGACAAATATTGACCAAGACATTAATA 1852
 Db 12529 GCAGATGCTTGAAGATGCCCCAAATCTGACAGAGACAAATATTGACCAAGACATTAATA 12588
 QY 1853 CTTGAAAAAATAAAAAAAAAAAAAAAAAAANNC 1887
 Db 12589 CTTGAAAAAATAAAAAAAAAAAAAAAAAAANNC 12623

RESULT 4
 ADM02768
 ID ADM02768 standard; cDNA; 3307 BP.
 XX AC
 XX ADM02768;
 XX AD
 XX 20-MAY-2004 (first entry)
 XX DE
 XX Human cDNA of the invention SEQ ID NO:1453.
 XX KW
 XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
 XX OS
 XX Homo sapiens.
 XX PN
 XX EP1347046-A1.
 XX PD
 XX 24-SEP-2003.
 XX FP
 XX 12-APR-2002; 2002EP-00008400.
 XX PR
 XX 22-MAR-2002; 2002JP-00137785.
 XX PA
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX PI
 XX Iobagai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 XX WPI: 2003-723558/69.
 XX DR
 XX P-RSDB: ADM05211.
 XX PT
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 XX developing a diagnostic marker or medicines for regulating their
 XX expression and activity, or as a target of gene therapy.
 XX Claim 1; SEQ ID NO 1453; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM6202-ADM6773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM0316-ADM0358 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.

SQ Sequence 3307 BP; 1122 A; 632 C; 829 G; 724 T; 0 U; 0 Other;

Query Match 80.6%; Score 1637.6; DB 11; Length 3307;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;

QY 8 GCTGACAGAAACAAATCAATCTCTCTCAGC-TCCCGCCCCAGCCCTTGAAATATGATAC 66
DB 1457 GCTGACAGAAACAAATCAATCTCTCTCAGCTTCCCGCCCCAGCCCTTGAAATATGAAAC 1516
QY 67 TCTAAGCCAGCAGCAGAGAAACATGCGCACTGCGTGAATTGATAGCTGAACAAGCC 126
DB 1517 TCTAAGCCAGCAGCAGAGAAACATGCGCACTGCGTGAATTGATAGCTGAACAAGCC 1576
QY 127 TCAATATAGATTAAGTGAACAAAAGCTGG-CCACAGTACTGAATGAGCCCTGGGGAAG 182
DB 1577 TCAATATAGATTAAGTGAACAAAAGCTGGCCACAGTACTGAATGAGCCCTGGGGAAG 1636
QY 183 CTTTCTATCCAGAGAGAGATGTGCGAGCCGACACCCCTTACAGTCACTTAAGAGAG 242
DB 1637 CTTTCTATCCAGAGAGAGATGTGCGAGCCGACACCCCTTACAGTCACTTAAGAGAG 1696
QY 243 TGTCAAAAAAGCGTGTGCGCACTGATGAAGCCATTTTCATCACTGATGTTCCATGA 302
DB 1697 TGTCAAAAAAGCGTGTGCGCACTGATGAAGCCATTTTCATCACTGATGTTCCATGA 1756
QY 303 CAAATATAGATCAATCTTGAGAGCCGTGAAGCATGTGGAAGCTGTGAGGAGCCACC 362
DB 1757 CAAATATAGATCAATCTTGAGAGCCGTGAAGCATGTGGAAGCTGTGAGGAGCCACC 1816
QY 363 CTCTATCTCTGAGAGGTTGAGAAATCAAGAAACAGATCACTGAATAATGAATGTGTC 422
DB 1817 CTCTATCTCTGAGAGGTTGAGAGATCAAGAAACAGATCACTGAATAATGAATGTGTC 1876
QY 423 AGTAGACATGGAAGAAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGAAAT 482
DB 1877 AGTAGACATGGAAGAAAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGAAAT 1936
QY 483 GATTGCTAGATCTGGGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTGAGATTAAGCT 542
DB 1937 GATTGCTAGATCTGGGGGGGAGCTGATTAAGACATATCTGCCAAAGCTGTTGAGATTAAGCT 1996
QY 543 TGAACCAATGCTTTTCAATTTGGAGAACATACACACTGCTGGAAGAGAGAGGCCAA 602
DB 1997 TGAACCAATGCTTTTCAATTTGGAGAACATACACACTGCTGGAAGAGAGAGGCCAA 2056
QY 603 ACTACTGATGTATGAGAGCTACAGAAAGTTCTGTGTGATCACTGTCATTGATAGT 662
DB 2057 ACTACTGATGTATGAGAGCTACAGAAAGTTCTGTGTGATCACTGTCATTGATAGT 2116
QY 663 TACCATTTAAGATACTCAAGATTTCATCCGGGAGCTGGAAGAGCTGGAATGATGCTTC 722
DB 2117 TACCATTTAAGATACTCAAGATTTCATCCGGGAGCTGGAAGAGCTGGAATGATGCTTC 2176
QY 723 AGTAGTAAACCAACAGCAAGAACAGCAGACCATTAAGGGAAGAAATAGATGACTACA 782
DB 2177 AGTAGTAAACCAACAGCAAGAACAGCAGACCATTAAGGGAAGAAATAGATGACTACA 2236
QY 783 GGAAGAGCTGAGATATGTTATTAACTGATGTTGAACTCATTTGGCGCATGTGGGAGCC 842
DB 2237 GGAAGAGCTGAGATATGTTATTAACTGATGTTGAACTCATTTGGCGCATGTGGGAGCC 2296

QY 843 TGATAAACCCATTGTCAAGAAAGATATACATGATTAATTCAGCATGGGATTTCTTAA 902
DB 2297 TGATAAACCCATTGTCAAGAAAGATATACATGATTAATTCAGC-ATGGGATTTCTTAA 2355
QY 903 ATTAAGCTTGAAGAAAGCCGATTTGACCAATTGAG-AGCAATGACGTCCGCTCA 956
DB 2356 ATTAAGCTTGAAGAAAGCCGATTTGACCAATTGAGGAGCAATGACGAGCTGCGCTTCA 2415
QY 957 GTACAGATGAGCTGCGAGGCGGATTTTTTTCTGGGTATATTTGAGGTGTTAGTTGCTT 1016
DB 2416 ACCAGATGAGCTGCGAGGCGGATTTTGACTGGGTATATTTGAGGTGTTAGTTGCTT 2475
QY 1017 CAATGCTCCAAATTTGGAACAGCATCTCGAAACGTCAACAGAGAGATTTGATGCTTAGAG 1076
DB 2476 CAATGCTCCAAATTTGGAACAG-ATCTCGAAACGTCAACAGAGAGATTTGAGGCTTAAG 2534
QY 1077 CAATTTAAGCTTGAGGCTTATCAACAGAGATAG-AGTGTACGACTGACTATCAAGCA 1135
DB 2535 CAATTTAAGCTTGAGGCTTATCAACAGAGATAGGAATGGAAGAAAGACTGATCATCAAGCA 2594
QY 1136 GACGCTTTTGTCTAAAGAAAGTAAACAGAGAGAGTGAACAAACACTGTTCTAGACCCAA 1195
DB 2595 GA-GCTTTTGTCTAAAGAAAGTAAACAGAGAGAGTGAACAAACACTGTTCAAGACCATT 2653
QY 1196 AATGAGCTGATTAATGATAT-GGCTAGCCCTGGAGNAGAGATCATCAACAGACGATTA 1254
DB 2654 AATGAGCTGATTAATGATATGAGATAGCTTGAGAGAGAAATCATCAACAGACGATTA 2713
QY 1255 ACTGAGAGGTGCTATTAAGCCNTGGGTCAAGTCCCAACATGCCCTGATAGGCTCTG- 1312
DB 2714 ACTGAGAGGTGCTATTAAGCTTGGGTCAAGTCCCAACATGCCCTGATAGGCTCTG 2773
QY 1313 CATGCTGACACACACGAGGCTTGTAGTAGAGCAAGAAACCTGTTGAGAGAGACCTTA 1372
DB 2774 ATGGCTGACACACACGAGGCTTGTAGTAGAGCAAGAAACCTGTTGAGAGAGACCTTA 2833
QY 1373 AGCCATTGAATTTGAACCTTGCAAGCATGTGCTCCAAATATGATGTTATGCCCCATCA 1432
DB 2834 AGCCATTGAATTTGAACCTTGCAAGCATGTGCTCCAAATATGATGTTATGCCCCATCA 2893
QY 1433 GTTCCACATGGAAGCCGTTAATTAAGACAGAAATGATCTTAATGTAATCAAGTGCAGAGA 1492
DB 2894 GTTCCACATGGAAGCCGTTAATTAAGACAGAAATGATCTTAATGTAATCAAGTGCAGAGA 2953
QY 1493 AGAAGCAACCACTTCAAGAACAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTT 1552
DB 2954 AGAAGCAACCACTTCAAGAACAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTT 3013
QY 1553 GGAAGAAACAGAACAAAGAGAGAGAGCTGATGAGTGTGCTTGGCCAGAGCCAAAGGTT 1612
DB 3014 GGAAGAAACAGAACAAAGAGAGAGAGCTGATGAGTGTGCTTGGCCAGAGCCAAAGGTT 3073
QY 1613 CCATGGCGAAATTTGAGATTTGACAGAGTGGCTGACTGACAGAGAGCTCATCTGTTGCG 1672
DB 3074 CCATGGCGAAATTTGAGATTTGACAGAGTGGCTGACTGACAGAGAGCTCATCTGTTGCG 3133
QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGTCATATGGA 1732
DB 3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGTCATATGGA 3193
QY 1733 AGTGTGTGCTGCTTTGAAGCTTAAGAAAGAAACATTAAGAGCTGATCAGAAAGGCA 1792
DB 3194 AGTGTGTGCTGCTTTGAAGCTTAAGAAAGAAACATTAAGAGCTGATCAGAAAGGCA 3253
QY 1793 GCAGATGCTTGCAGAGTGCCTCAAAATCTGCAGAGACAAATATTTGACCAAGACAT 1846
DB 3254 GCAGATGCTTGCAGAGTGCCTCAAAATCTGCAGAGACAAATATTTGACCAAGACAT 3307

RESULT 5
AAH41086
ID AAH41086 standard; cDNA; 6006 BP.

XX	AA#41086;	
AC		
DT	29-AUG-2001	(first entry)
XX		
DE	Human calcium binding/actin crosslinking protein (CBACP) cDNA.	
XX		
KM	Human; calcium-binding/actin cross-linking protein; CBACP; infertility;	
KM	immunological disease; hypertension; renal insufficiency; infection;	
KM	muscular dystrophy; malnutrition; neurotransmission disorder; cardiac;	
KM	kidney disease; anaemia; heart disease; cancer; immunomodulatory;	
KM	hypertensive; nephritic; antiviral; antibacterial; antianemic;	
KM	cytostatic; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FH	CDS	1147..4992
FT		/tag= a
FT		/product= "CBACP"
FT		/note= "Calcium binding/actin crosslinking protein"
XX		
PN	W0200138391-A1.	
XX		
PD	31-MAY-2001.	
XX		
PE	24-NOV-2000; 2000WO-CN000492.	
XX		
PR	24-NOV-1999; 99CN-00124335.	
XX		
PA	(SHAN-) SHANGHAI BIORIGIN GENE DEV CO LTD.	
XX		
P1	Mao Y, Xie Y;	
XX		
DR	WPI: 2001-367664/38.	
XX	P-PSDB; AAG62365.	
PT	Calcium-binding/actin cross-linking protein and encoded polynucleotide,	
XX	applicable in diagnosis and treatment of diseases due to abnormality of	
PT	the polynucleotide and polypeptide including infertility and	
XX	hypertension.	
XX		
XX	Claim 6; Page 37-40; 48pp; Chinese.	
XX		
CC	This invention relates to human calcium-binding/actin cross-linking	
CC	protein (CBACP) and its encoding cDNA sequence. Included in the invention	
CC	are vector containing the CBACP nucleotide sequence, a host cell	
CC	transformed with the vector and an antibody directed against the CBACP	
CC	protein specifically. The protein and its encoding nucleotide sequence	
CC	can be used in the diagnosis and treatment of infertility, immunological	
CC	diseases, hypertension, renal insufficiency, muscular dystrophy,	
CC	malnutrition, neurotransmission disorder, viral and bacterial infections,	
CC	kidney diseases, anaemia, heart diseases and cancer. Use of the protein	
CC	or nucleotide sequence may result in immunomodulatory; hypotensive;	
CC	nephritic; antiviral; antibacterial; cardiac; antianemic or cytostatic	
CC	activity. The present sequence represents cDNA encoding human CBACP	
XX		
SQ	Sequence 6006 BP; 1867 A; 1249 C; 1471 G; 1419 T; 0 U; 0 Other;	
XX		
Query Match	49.7%;	Score 1010; DB 4; Length 6006;
Best Local Similarity	93.9%;	Pred. No. 1.5e-234;
Matches 1126; Conservative	0;	Mismatches 60; Indels 13; Gaps 7;
QY	699	GGAGATCCTGGATTTGATTCCTTCAGTGTAAACACAGCAAGAGAGAGACAT 758
DB	903	GGAGATCCTGGATTTGATTCCTTCAGTGTAAACACAGCAAGAGAGAGACAT 962
QY	759	AAGGAGAAATATGATGGACTACAGAGAGAGCTGGATATAGTTATTAACTAGATTCTGA 818
DB	963	AAGGAGAAATATGATGGACTACAGAGAGAGCTGGATATAGTTATTAACTAGATTCTGA 1022
QY	819	ACTCATTTGGGCGATGTGGGAGCCTGATTAACCATTTGTCAAGAGATATACATGAGTT 878

Dd	1023	ACTCATTTGGCGCAATGTTGGGAGAGCCGATTTAAACCATTTGTCAAGAAAGATATAGTAGTT	1082
Oy	879	AAATTCAGCGATGGGATTTCTTAATATAAGCTTGGAAAGACCGGATGGACCACTTAG- 937	
Dd	1083	AAATTCAGC-ATGGGATTTCTTAATATAAGCTTGGAAAGACCGGATGGACCAACTTAGG 1141	
Oy	938	-----AGCAATGCAAGCTGCCGTCAAGTACAGATGGACTGCGAGCGGTATTTTTCTGGGTAG 992	
Dd	1142	AGGCAATGCAAGGCTGCCGTTCAGTACCAAGATGCACTGACAGGCGGTATTTTGACCTGGGTAG 1201	
Oy	993	ATATTTGCAAGGTGTAAGTTCGCTTCAATGTCTCCAAATTGGAAACAGCAATCTCGAAACTGTC 1052	
Dd	1202	ATATTTGCAAGGTGTAATTTAGCTTTCAATGTCTCCAAATTGGAAACAG-ATCTCGAAACTGTC 1260	
Oy	1053	AAGCAGCAGATTTGATGTGCTAGAACCAATTTAAAGCTGAGGCGCTATCAACAGCAGATAG-A 1111	
Dd	1261	AAGCAGCAGATTTGAAGAGCTTAAAGCAATTTTAAAGCTGAGGCGCTATCAACAGCAGATAGAA 1320	
Oy	1112	GTGGTAGACTGATCTCATCACGCAAGCGCTTTTGCTTAAAGAAAGTAAACAGNAGAGATGA 1171	
Dd	1321	ATGGAAGAGACTGAATCTATCAAGCAGA-GCTTTTGCTTAAAGAAAGTAAACAGNAGAGATGA 1379	
Oy	1172	CAAAACACACTGTTTAGACCCCAAAAATGAGNCTGATTTTGATAT-GGCTAGCGCTGAGNA 1230	
Dd	1380	CAAAACACACTGTTTAAAGACCATTTAAAGAACTGAAATTTGATATGAGATAGCTTGGAGGA 1439	
Oy	1231	GAGAAATCATCAACAGACGATAAACTGGAAGGTGCTCTATTAGCCNATGGGTCACTTCCA 1290	
Dd	1440	GAGAAATCATCAACAGACGATAAACTGGAAGGTGCTCTATTAGCTTTGGGTCACTTCCA 1499	
Oy	1291	ACATGCCCTGTATAGCTCCTG--CATGCTGACACACACCGAGGGCTTGCTTAAGTGACA 1348	
Dd	1500	ACATGCCCTGTATAGCTCCTGCGCATGGCTGACACACACGAGGGCTTGCTTAAGTGACA 1559	
Oy	1349	GAAACCTGTTGAGAGAAACCTTAAAGCAATTGAAATTGAACTTGCCAAAGCATCATGAGCT 1408	
Dd	1560	GAAACCTGTTGAGAGAAACCTTAAAGCAATTGAAATTGAACTTGCCAAAGCATCATGAGCT 1619	
Oy	1409	CCAAATATGATGTATTAGCCCATCAGTCCACAGTGGAAAGCCGTTTAAATAAGCAGGAAATGA 1468	
Dd	1620	CCAAATATGATGTATTAGCCCATCAGTCCACAGTGGAAAGCCGTTTAAATAAGCAGGAAATGA 1679	
Oy	1469	TCTAATTGAATCAAGTGCAGAGAGAAAGACAGCACTTTCAGAACAACTAGAGTTTTT 1528	
Dd	1680	TCTAATTGAATCAAGTGCAGAGAGAAAGACAGCACTTTCAGAACAACTAGAGTTTTT 1739	
Oy	1529	AAATCAACGCTGCGAAATGTTTTGGAAAAAACGAACAAAGGAAGACAGCCTGATGAG 1588	
Dd	1740	AAATCAACGCTGCGAAATGTTTTGGAAAAAACGAACAAAGGAAGACAGCCTGATGAG 1799	
Oy	1589	TGCTTTCGCGACGAGCCAAAGGGTTCCATGCGGAAATTTGAGAGATTTTGCAGCAGTGGCTGAC 1648	
Dd	1800	TGCTTTCGCGACGAGCCAAAGGGTTCCATGCGGAAATTTGAGAGATTTTGCAGCAGTGGCTGAC 1859	
Oy	1649	TGACACGAGAGCGTATCTGTTGGCATTTAAACCGCTGGAGGTTTACCGGAAACAGCCAA 1708	
Dd	1860	TGACACGAGAGCGTATCTGTTGGCATTTAAACCGCGGAGGTTTACCGGAAACAGCCAA 1919	
Oy	1709	GGAGCAGCTTAATGTCCATATGAAAGTGTGTGCTTCTTGAAGCTTAAAGAAACATTA 1768	
Dd	1920	GGAGCAGCTTAATGTCCATATGAAAGTGTGTGCTTCTTGAAGCTTAAAGAAACATTA 1979	
Oy	1769	TAAAGAGTCTGATGCAAGAAAGGCGCAGCAGATGCTTGCAGAAAGTCCCAAAATCTGCAAGAC 1828	
Dd	1980	TAAAGAGTCTGATGCAAGAAAGGCGCAGCAGATGCTTGCAGAAAGTCCCAAAATCTGCAAGAC 2039	
Oy	1829	AAATATTTGACCAAGACATAAATTAACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAANAC 1887	
Dd	2040	AAATATTTGACCAAGACATAAATTAACCTTGAAGAAATTTGGAAATCGGTGAAACCAAAAC 2098	

ID	ACAA56949 standard; cDNA, 6007 BP.
XX	
AC	ACAA56949;
XX	
DT	10-JUN-2003 (first entry)
XX	
DE	Human adipocyte Selected Interacting domain, SID, cDNA #36.
KM	Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
KM	anorectic; antidiabetic; protein-protein interaction; diabetes;
XX	Yeast 2-hybrid assay; metabolic disorder; obesity.
XX	
OS	Homo sapiens.
XX	
PN	WO200286122-A2.
PD	31-OCT-2002.
XX	
PF	14-MAR-2002; 2002WO-EP003768.
XX	
PR	14-MAR-2001; 2001US-0275734P.
PA	(HYBR-) HYBRIGENICS.
PI	Legrain P, Daviet L;
DR	WPI; 2003-103412/09.
XX	
DR	P-PSDB; AEU70405.
PT	New complex between two interacting proteins in adipocyte cells, useful
PT	for identifying selected interacting domains that modulate protein
PT	interactions, or for preventing or treating metabolic disorders such as
XX	obesity or diabetes.
PS	
PS	Claim 7; Page 120-122; 382pp; English.
XX	
CC	The invention relates to a complex between two interacting proteins in
CC	adipocyte cells, given in the specification. The proteins are identified
CC	by selecting a bait protein from a known adipocyte marker and then
CC	performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC	members of an adipocyte cDNA library. The proteins are designated SID
CC	(RTM) (selected interacting domains) proteins. Also included are a
CC	polynucleotide encoding a polypeptide in the adipocyte cells, a
CC	recombinant host cell expressing at least one of the interacting
CC	polypeptides of the complex, selecting a modulating compound in adipocyte
CC	cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC	sequences given in the specification (including its fragment or variant),
CC	a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC	given in the specification (including its fragment or variant), a vector
CC	comprising the SID (RTM) polynucleotide, a recombinant host cell
CC	comprising the vector, a protein chip comprising the polypeptides and a
CC	record comprising all or part of the data, listed in the specification.
CC	The complex, polypeptides, polynucleotides and compounds are useful for
CC	preventing or treating metabolic disorders such as obesity or diabetes.
CC	The polynucleotides are useful as probes or primers. The complex is
CC	particularly useful for identifying selected interacting domains (SID
CC	(RTM)) for screening drugs that modulate the protein interaction, thus
CC	exhibiting the therapeutic effect. The present sequence encodes a SID
CC	(prey) protein of the invention
XX	
XX	
SO	Sequence 6007 BP, 1810 A; 1357 C; 1550 G; 1290 T; 0 U; 0 Other;
	Query Match 32.3%; Score 655.4; DB 8; Length 6007;
	Best Local Similarity 62.9%; Pred. No. 1.8e-148;
	Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
QY	
DB	
1259	CCCCGCCAGCCCTTGAATATGATATCTTAAAGCAGCAGCAGGAAGAAATCGGCAACTG 100
1259	CCCTCTCCAGCATGATGATCATGAGCAGCTCAGCGAGCAACAAGAGGAATATGAGCAATTA 1318
101	CGTAGCTGATAGTGCAGACACAAAGCCTCATATTAATTAAGATGAACAAACTGGCC---A 156
1319	AGGGATCTTATTTGCGAACACAAACTCATATTTGCAAACTATTAAGATGAGGCCACAA 1378

OY	157	CAGTATCTGAATGACACCCCTGGGGAGAGCCCTTTTCTATCCAAAGAGATATGTGTGACGCCGAC	216
Db	1379	CTAAAGGAATTTAAACCCCTGAGAGAAAGGGGAAATGGTGGAAAGAAAAATATCCAAAGAGCGAAA	1438
OY	217	ACCCTTTACAGTGCACATTTAAAGAGATGTCAAAAAGCGTGTGTGGCACTGGATGAAAGCC	276
Db	1439	AACATGTATGGCCAAATTAAGAGAGAGAGTGGCCAGAGAGCCCTGGCTCTGGATGAAAGCC	1498
OY	277	ATTTCCTCAATCAATC-----AGTTCCATGACAAATAGATCAGATCCTTGGAGAC	327
Db	1499	GTGTCCCAAGTCCACACAGATTAACAGAGTTCATATATAAATTGAGCCCTATGTGTGAGACA	1558
OY	328	CTGGAGCGCATCGTGGAAACGCTGTAGGAGAGCACCCCTCTATCTGTGACAGAGTTGAGAG	387
Db	1559	CTGGAGAAATCTTTCCTCTCGCTTCGCTGTATGTCCACACACTGATTCCTGCTGAAATGACAG	1618
OY	388	ATCAAGGACAGATCAGTGGAAAAATTAAGATGTGTCAGTAGACATGAAAGAGCTACAGCCG	447
Db	1619	ATCAGAGAGTGCATCAGTGAACAAATAAGAGTCCACCGTGAAGCTAAGAAAACTGCAGCCA	1678
OY	448	TTGTATGAATCTCTTAAACAGAGGGGAGAGAAAATGATTCGTAAATCTGGGGGACCTGAT	507
Db	1679	TCTTTTGAAGGCTTTGAAGAGCGCCGTGAGAGAGAGCTTATGTGACATCTCAGGGAGCAGAC	1738
OY	508	AAAGACATATCTGCCCAAGGCTGTCAGAGATTAAGCTTAACCAATAGTGTTCATTTTGGGAG	567
Db	1739	AAGGATCTGGCTGCAAAAAGAAATCCAGAGATTAATTTGATCAAAATGTATTTCTTCTGGGAG	1798
OY	568	AACATACACACACTGTGTGGAGAGAGAGGAAAGCCAACTACTGTGATGTGTAGAGACTACGA	627
Db	1799	GACATCAAAAGCTCGGGGCTGAAGAAACGAGAAATCAAAATTTCTTGATGTGTCCTTGAATTTGACA	1858
OY	628	GAAAGTTCTGTGTGTATACATGTCAATTGATAGTTACATTAAGATTAATCAAGATTTTC	687
Db	1859	GAGAAATTCTGTGTATGACATGTGACAGCTCTCTGACACACATCAAAAGACACCCAGAGATATT	1918
OY	688	ATCCGGGACCTGGAAAGATCTCTGAATTGATCTTCAGTAGTAAACACAGCAAGAAAGCA	747
Db	1919	GTCCATATCACTTGGAAAGCCCGACGATTAATCTTTCATCATCAAAACAAACAGTTGAAGACT	1978
OY	748	GCAGAGACCAATTAAGGAGAAATATGATGTGATCAAGAGAGGCTGGATATGATTAATTAAC	807
Db	1979	GCTGAGACTATTAAGAGAGAGACAGATGTGTGATGATAGCTGTGGAGTTTATTCGGATTC	2038
OY	808	CTAGGTTCTGAACATCTTGGCGGCATGTGTGGGAGCCTGATTAACCCATTTGTCMAAGAGT	867
Db	2039	CTTGAGACAGATTTGATTTTGTTCCTGTGTGAGAAACTGAGAAAGCCTGAAGTGAAGAGAGC	2098
OY	868	ATACATAGATTAAATTCAGCGATGTGGATTTCTTAATTAAGCTTGGAAAGACCGGATTTGA	927
Db	2099	ATTGATAGATGAATTAATGTC-TTGGAGAACTTTAAACAAACATGAGAAAGAGAGGCTAGA	2157
OY	928	CCAACCTTGAG-----AGCAATGTGAGCTGCGCTGAGTACAGATGAGACTGAGCGGGTATT	981
Db	2158	AAAACCTTGAGAGATCTATGCAAGCTGCTGTGTGACATATCAGAGACACTTTCAAGCTATGTT	2217
OY	982	TTTCTGTGGTAGATATTGCAAGGTGTAAATGTCCTTAATGTCTTCAAATTGGAACAGCATC	1041
Db	2218	TGACTGTGCTAGATTAACCTGTGATTAACCTCTGACCATGCCCCCTGTGGCACTG-ACC	2276
OY	1042	TTCGAATCTGTCAACACAGACAGATTGATGTGTGAGAGCAATTAAGTGTGAGAGCCATATAC	1101
Db	2277	TCAATTAATGTTTAAAGATCAAGTTAAATTAATGAAGAGATTCAAGTAGAAGTTTACCAAC	2336
OY	1102	AGCAGATAGATGTGTACGACTGACTCATCAACGACAGCGCTTTTGTCTTAAGAAATGAACG	1161
Db	2337	AGCAAAATTGAGATGAGAGAGCTTAATATCAACAGGGGTGAACGTATGTTAAAGAAATCTACTG	2396
OY	1162	NAGAGATGACAAACACACTGTTCTTAACCCCAAAAATGAGATGTATATTGATAT-GGCTA	1220
Db	2397	ATGAGACGGACAGAGACATTATACAGAAACCACTGACAGAACTTCAACACCTCTGGGAGA	2456

QY 1221 GCCTGAGNAGATCATCAACAGACAGATAAAGGAGGCTCTATTAAGCCTG 1280
 DB 2457 ACCCTGGGAGAAAATTGCCACGACGACAACTGAAAGGGCTCTGTGGCCCTTG 2516
 QY 1281 GTGAGTTCCACATGCCCCCTGTATAGCTCTGCAT--GCTGACACACACGAGGCTTGC 1338
 DB 2517 GTGAGTTCCAGCATGCTTATAGAGAACTAATGAGTTGGCTGCTACATCCGAAAGGTTGT 2576
 QY 1339 TAAATGACGACAAACCTGTGTGAGAGACCTTAAAGCATTGAATTAATGACTTGCACAGC 1398
 DB 2577 TAGATGCTCAGAGACCAATAAGTAGAGAGCCAAAAGTATTAAGTTGAGCTCCCAAGC 2636
 QY 1399 ATCATGTGCTCCAAATGATGATATTAAGCCATCGATCCAGTAGAGAGCCGTTATATAAG 1458
 DB 2637 ACCATGTCCTTAAAAATGATGATTTTGGCTCATCAAGCCACAGTGAACAGTCAACAAAG 2636
 QY 1459 CAGGAAATGATCTAATTAATCAAGTGCAGAGAGAAAGCAAGCAACCTTGACAAACAGC 1518
 DB 2697 CTGGCAATGAGCTTCTGAATCCAGTGTGAGATGATGCCAGCAGCTTAAGAGCCGTT 2756
 QY 1519 TAGAGGTTTAAATCAACGCTGCAAAATGTTTGGAAAAACAGAACAAAGAACAGC 1578
 DB 2757 TGGAAAGCCATGAACCAATGCTGGAGAGTCAGTGTTAACAGAAAAACAGAGAGAGGAGCAGC 2816
 QY 1579 AGCTGATGCTGCTGCTGGCCAGGCAAAAGGTTCCATGAGGAAATTGAGGATTTGACAGC 1638
 DB 2817 AGCTTCAGTCACTCTGACAGAGGCCAGGCTTCCACAGTAATGAATGAAATTTCTCT 2876
 QY 1639 AGTGGCTGACTGACAGGAGCGTCATCTGTGGCATCTTAAACCGCTGGAGAGTTTACCGG 1698
 DB 2877 TGGAACTTACTAGAAATGAGAGCCAGCTTTCTGATCTTAACCCACAGAGAGACTTCTCG 2936
 QY 1699 AAACAGCCAGAGAGCAAGCTTAATGTCATATGAGAACTGTGCTGCTCTTGAAGCTAAAG 1758
 DB 2937 AAATCTGTAGGGAAACAGCTGTATACATATGAGAACTTATCCAGCTGAAAGCCAAAG 2996
 QY 1759 AAGAACTAATTAAGTCTGATGAGAAAGGCCAGACAGATCTTGCAAGATGCCAAAT 1818
 DB 2997 AAGAGACTTAATTAACACTACTTGAACAGGAGACATCATCTTCTTAAGCCGTAGCAGCT 3056
 QY 1819 CTG 1821
 DB 3057 CTG 3059
 RESULT 7
 AAS22908
 ID AAS22908 standard; DNA; 16122 BP.
 XX
 AC AAS22908;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE DNA encoding novel bone marrow polypeptide #2.
 XX
 KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
 KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
 KW wound healing; nutritional supplement; immune disorder;
 KW severe combined immunodeficiency; SCID; ds.
 XX
 OS Homo sapiens.
 XX
 FN MO200157187-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003782.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 30-NOV-2000; 2000US-0250683P.
 XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Ford JB, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 PI Ren F, Drmanac RT;
 XX
 DR WPI; 2001-488875/53.
 DR P-PSDB; AAU14603.
 XX
 PT Nucleic acid encoding bone marrow polypeptides, useful in diagnostic and
 PT gene therapy.
 XX
 PS Claim 1; Page 146-151; 392pp; English.
 XX
 CC AAS22907-AAS23099 represent nucleic acids encoding novel bone marrow
 CC polypeptides. The nucleic acids and corresponding proteins may be used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate bone marrow polypeptide expression. For example, to treat
 CC disorders associated with decreased expression by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of the
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of the polypeptide. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic
 CC acids into a host cell and culturing the cell to express the protein. The
 CC nucleic acid and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and therefore which patients
 CC may be in need of restorative therapy. The proteins may also be used as
 CC antigens in the production of antibodies against bone marrow proteins and
 CC in assays to identify modulators of their expression and activity. The
 CC anti-bone marrow protein antibodies and antagonists may also be used to
 CC down regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
 CC used to regulate haematopoiesis activity, and consequently in the
 CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
 CC as wound healing; as a nutritional supplement; and in treatment of
 CC immune disorders such as severe combined immunodeficiency (SCID)
 XX
 SQ Sequence 16122 BP; 4917 A; 3572 C; 4169 G; 3464 T; 0 U; 0 Other;
 Query Match 32.3%; Score 655.4; DB 4; Length 16122;
 Best Local Similarity 62.9%; Pred. No. 2.6e-148;
 Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
 QY 41 CCCGCCCCAGCCCTTGAATATGATCTTAAAGCAGACAGGAAAGAAATCGGCACTG 100
 DB 11377 CCCCTCCAGCATTAATGATGACAGCTCAGCAGCAACAAAGAGGAAATGAGCAATTA 11436
 QY 101 CGTAGTTGATAGCTGAACCAACGCTCATTAATGAATGAACAAACTGGCC---A 156
 DB 11437 AGGGAATCTAATGCTGAACCAAACTCATATTAACAACTAATGAATGAGCCACAA 11496
 QY 157 CAGTACTGAATGAGCCCTGGGGAAGGCTTTCTATCCAGAGAGATATGTGGACCGGAC 216
 DB 11497 CTAAGAATTAATTAACCTGAGAGAGGGAATGTGAGAAAGAAATACAGAAAGCAGAA 11556
 QY 217 ACCCTTACAGTCAATTAAGAAGATGTCAAAAAGCTGTGTGSCATGATGAAGCC 276
 DB 11557 AACATGTATGCCCAATTAAGAGAGGTGGCCAGGAGGCCCTGTGATGAAGAGCC 11616
 QY 277 ATTTCATCAATCAATC-----AGTTCATACAGATAGATCAAGATCTTGAAGAC 327
 DB 11617 GTGTCCAGTCCACAGATTAACAGAAATTTCTATTAATTAATGAGCCATGTTTGAAGCA 11676
 QY 328 CTGGAAGCATGCTGTAAGCTGTAGGACAGCAGCCCTCTATCTTGCAGAGAGTTGAAG 387
 DB 11677 CTGGAAGATCTTCTCTCTGCTGCTGCTGATGACACATGATCCCTGCTGAAGTGAAG 11736
 QY 388 ATCAAGAACAGATCAGTGAATTAAGAAATGTGTCACTAGACATGAAGAAAGCTACAGCC 447
 DB 11737 ATCAGAGAGTGCATCAATTAAGATGAGCCACCGTGAAGCTGAAGAAAGCTCAGAGC 11796
 QY 448 TTGTATGAATCTTAAACAGAGGAGAGGAAATGATTTGCTAATCTGGGGGAGACTGAT 507

Db	11797	TCCTTTGAGGCTTGAAAGCCCGCTGGAGAGAGCTTAATTGACGATCTCAGGAGGACAGAC	11856
Qy	508	AAAGACATATCTGCCAAAGCTGTTCAAGATTAAGCTTGACCAAAATGTTTTTCATTTGGGAG	567
Db	11857	AAGGATCTGGCTGCAAAAAGAAATCCAGAGTAATAATGGAATCAAAATGGTATTTCTTCGGAG	11916
Qy	568	AACATACACACACTGTGTGGAAAGAGAGGAAAGCCAACTACTGATGTGATGAGCTAGCA	627
Db	11917	GACATCAAAAGCTCGGGCTGGAAAGAAACGAAATCAAAATTTCTGTATGTCTTGAATTTAGCA	11976
Qy	628	GAAAAGTTCTGATGATCACAATGTCATGTATGATTAACCTTAAGATATCTCAAGATTTC	687
Db	11977	GAGAGTTTCTGTATGTACATGCGACGCTCTCCTGACCACTCAAGACACCCGAGATATT	12036
Qy	688	ATCCGGGACCTGGAAAGATCCTGGAATTGATCTTCACTAGTATTAATAACACAGCAAGAGCA	747
Db	12037	GTCCATGACTTGGAAAGCCAGGAGTATGATCTTCCATCATCAAAACAAGGTGGAAGCT	12096
Qy	748	GCAGAGACCATTAAGGGAAGAAATATGATGACTACAGAGAGAGCTGATATATTAATTAAC	807
Db	12097	GCTGAGACTATTAAGGAAGACAGATGATCTGCATGATGAAGCTGAGGATTATTCGATTC	12156
Qy	808	CTAGGTTCTGAACATTTGGCGGCATGTGTGGGAGCTTGATTAACCCATGTGCAGAAAGAT	867
Db	12157	CTGGAGCAGATTTGATTTTTCCTGTGGAGAAACTGAGAAAGCTGAAGTGAAGAAAGAC	12216
Qy	868	ATACATGAGTTAAATTCAGCGATGGATTTCTTAATTAAGTCTGGAAGAACCGGATTTGA	927
Db	12217	ATTGATGAGATGAATTAATGTC-TTGGAGAACTTTAAACAAACATGGAAGAGGCTTA	12275
Qy	928	CCAACTTAG-----AGCAATGCACTGCCCTGCATGACAGATGACTGACGGCGGATTT	981
Db	12276	AAAACTTGAGATGCTATGCAAGCTGCTGTGCACTATCAAGACACTTTCAGGCTATGTT	12335
Qy	982	TTTCGGGTAGATTAATTCGAGGTGTAAGTTTCGTTCAATGTCTCCATTTGGAAACAGATTC	1041
Db	12336	TGACTGCTAATATTAACCTGTGATTTAACTTCGACATGCCCCCTGTGTGGCACTG-AAC	12394
Qy	1042	TCGAAACTGTCAACAGACAGATTTGATGTGCTAGAGCAATTAAAGCTGAGGCTTATCAAC	1101
Db	12395	TCATATCTGTTAAAGATCAAGTTAATGTAATGAAAGAGTTCAAAAGTGAAGTTTACCAAC	12454
Qy	1102	AGCAGATTAAGATGTGACGATGACTCATCAAGCAGACGCTTTTGTAAAGAAATGAACAG	1161
Db	12455	AGCAAAATTGAATGAGAGAGAGCTTAATACACAGGGTGAACATGATTTAAAGAAAGCTACG	12514
Qy	1162	NAGAGATGAACAAACACACTGTTTTAACCCCAAAATGAGACTGATATTTGATTT-GGCTA	1220
Db	12515	ATGAGACGGAGAGAGACATTAATACAGAACCACTGACAGAACTTCAACACCTCTGGGAGA	12574
Qy	1221	GCTTGGAGNAGAGATCATCAACAGACAGATTAACCTGAGGGGTGCTTATTTAGACGNTGG	1280
Db	12575	AACCTGGTGAAGAAATTTGCCCAACGACACAAACTGAAGGGGCTCTGTGGCCCTTG	12634
Qy	1281	GTCAATTCAACATGCCCCGTGTATGAGCTCTGCAAT-GCTGACACAACCGAGGGCTTGC	1338
Db	12635	GTCAATTCAACATGCCCCGTGTATGAGAGCTTAATGAGTTGGCTGATCTATACCGAAGAGTTGT	12694
Qy	1339	TAAATGAGCAAAACCTGTGTGGAGAAACCTTAAGCCATTTGAATTTGAACCTTGCACAGC	1398
Db	12695	TAGATGTCCAGAGACCAATTAAGTGGAAACCCAAAGATCAATTTGAAGTTGAGCTCGCAAAAC	12754
Qy	1399	ATCATGTGCTCCAAATATGATTAATTTAGCCATCAGTCAACAGTGAAGGCGTTAATAAG	1458
Db	12755	ACCATGTCTTAAAAAATGATGTTTTTGGCTATCAACGCAACAGTGGAAACGTAACAAAG	12814
Qy	1459	CAGAAATGATCTAATTAATCAAGTCAAGGACAGAGAAAGCAACCTTCAAGAACAGC	1518
Db	12815	CTGGCAATGAGCTTCTTGAATTCAGTGTGGAAGTATGTCACACAGCTTTAAGGAGCCGTT	12874
Qy	1519	TAGAGTTTAAATCAACGCTGGCAAAATGTTTTGGAAAAACAGAACAAAGAGACAGC	1578

Db	12875	TGAAAGCCATGAAACCAATGCTGGGAGTCAAGTGTTTACAGAAACACAGAGGAGAGGACAC	12934
Qy	1579	AGCTGATGTGTGCTTCCTGGCCGACAGCCAAAGGATTCCAGTGGCGAAATTGAGATTGGCAGC	1638
Db	12935	AGCTTCAGTCAACTCTGTCAGACAGGCCAGGCTTCCACAGTGAATTTGAAGAATTTCCTCT	12994
Qy	1639	AGTGGCTGACTGACACGAGCGTCATCTGTGGCATCTTAAACCGCTGGGAGGTTTACCGG	1698
Db	12995	TGGAACCTTACTAGAAATGAGAGGCGACGCTTCTCATTCTTAAGCCACAGAGGAGCTTCC	13054
Qy	1699	AAACAGCCAGAGGACAGCTTAAATGTCCATPATGAGAGCTGTGCTGCTTTGAAGCTAAG	1758
Db	13055	AACTGCTAGGGAACAGCTTGAATACATATGGAACCTTAATTCACAGCTGAAGCCAGG	13114
Qy	1759	AAGAAACATATTAAGAGTCTGATGCAGAAAGGCCAGAGATGCTTGCAGATGCCCAAT	1818
Db	13115	AAGAGACTTATTAATCACTACTTGAACAAGGCGAGACTCATGCTTCAACCGTGAAGACT	13174
Qy	1819	CTG 1821	
Db	13175	CTG 13177	
RESULT 8			
AA523002			
ID	AA523002 standard; DNA, 16341 BP.		
XX	AA523002;		
AC			
XX	24-OCT-2001 (first entry)		
DT			
XX			
DE	DNA encoding novel bone marrow polypeptide #96.		
XX			
KW	Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;		
KW	haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;		
KW	wound healing; nutritional supplement; immune disorder;		
KW	severe combined immunodeficiency; SCID; ds.		
OS	Homo sapiens.		
XX			
PN	WO200157187-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US003782.		
XX			
PR	03-FEB-2000; 2000US-00496914.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	30-NOV-2000; 2000US-0250683P.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Ford JF, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;		
PI	Ren F, Dmanac RT;		
XX			
DR	WPI; 2001-488875/53.		
DR	P-PsDB; AAU14697.		
XX			
PT	Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and		
PT	gene therapy.		
XX			
PS	Claim 1; Page 253-257; 392pp; English.		
XX			
CC	AA523002-AA523099 represent nucleic acids encoding novel bone marrow		
CC	polypeptides. The nucleic acids and corresponding proteins may be used in		
CC	the prevention, diagnosis and treatment of diseases associated with		
CC	inappropriate bone marrow polypeptide expression. For example, to treat		
CC	disorders associated with decreased expression by rectifying mutations or		
CC	deletions in a patient's genome that affect the activity of the		
CC	polypeptides by expressing inactive proteins or to supplement the		
CC	patient's own production of the polypeptide. Additionally, the nucleic		
CC	acids may be used to produce the polypeptides, by inserting the nucleic		

CC acids into a host cell and culturing the cell to express the protein. The
CC nucleic acid and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins and
CC in assays to identify modulators of their expression and activity. The
CC anti-bone marrow protein antibodies and antagonists may also be used to
CC down regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
CC used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID)
XX
SQ Sequence 16341 BP; 4965 A; 3637 C; 4241 G; 3498 T; 0 U; 0 Other;
Query Match 32.3%; Score 655.4; DB 4; Length 16341;
Best Local Similarity 62.9%; Pred. No. 2.6e-148;
Matches 1134; Conservative 0; Mismatches 645; Indels 24; Gaps 7;
QY 41 CCCGCCCCGAGCCCTTGATATGATGATCTCTAAGCAGCAGCAGAGAAACATCGCACTG 100
DB 11596 CCTCTCCAGCATTTGATCATGAGCAGCTCAGGCAGCAACAAGAGAAATAGGCAATTA 11655
QY 101 CTTGATGATGATGCTGAACACAGCCTCATATGATTAAGTGAACAAACCTGGCC---A 156
DB 11656 AGGAAATCTATGCTGAAACAAACCTCATATTTGCAAACTAATAATAGGCCACAA 11715
QY 157 CAGTACTGATGAGCCCTGGGGAAGGCTTTCTATCCAGAAAGATATGTCGACCGAC 216
DB 11716 CTAAGGAATTAACCTCGAGAAAGGGGAAATGTTGAGAAATAATACCAAAAGCAAA 11775
QY 217 ACCCTTACAGTCACTTAATAAGAGATGTCAAAAACGCTGCTGTCGACTGATGAAGCC 276
DB 11776 AACATGATGCCCCAATTAAGAGAGAGTGCGCCAGCAGCCTCTGCTGATGAAGCC 11835
QY 277 ATTTCTCAATCACTC-----AGTTCAATGACAAAGATGATCAGATCCTTGAAGC 327
DB 11836 GTGTCCCACTCCACAGATTAACAGAGTTTCATGATTAATAATTGAGCCTATGTTGAGACA 11895
QY 328 CTGGAAGCCATCTGTGGAACCTCTGAGCAGCACCCTTATCTCTGCGAGAGTTGAAG 387
DB 11896 CTGGAAGATCTTCTCTCTGCTGCTGATGTCACCACTGATCCCTGCTGAAGTGAACAAG 11955
QY 388 ATCAAGGAACAGATCAGTGAATAATTAAGATGTGTCAATGATGATGGAAGAGCTACAGCCG 447
DB 11956 ATCAAGAGTGCATCATGAGACATTAAGAGTCCACCGTGAAGCTAGAAAACCTGACGCA 12015
QY 448 TTGTATGAACTCTTAAACAGAGGGGAGAGAAATGATTTGCTGATCTGGGGGGACTGAT 507
DB 12016 TCCCTTTGAGGCTTGAAGCGCGGTGAGAGAGCTTATTTGACATCTCAAGGAGACAGAC 12075
QY 508 AAACACATATCTGCCCAAGCTGTTCAAGATTAAGCTTGAACAAATGGTTTTCAATTTGGAG 567
DB 12076 AAGATCTGCTGCAAAAAGAAATCCAGATTAATGATTAATGATATTTCTTGGGAG 12135
QY 568 AACATACACACTGTGTGAAGAGAGAGCAAACTACTGATGTGATGAGAGCTACGA 627
DB 12136 GACATCAAAAGCTCGGGCTGAAGAGAGAAATCAAAATTTCTTGATGTCTTGAATTAGCA 12195
QY 628 GAAAAGTTCTGGTGTATCAGATGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 687
DB 12196 GAGAAAGTTCTGGTGTATCAGATGATGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 12255
QY 688 ATCCGGGAGCTGGAAGATCTGGAATTTGATCTTCAAGTAAAGTAAACACACAGAGAGCA 747
DB 12256 GTCCATGATCTTGAAGAGCCGAGCTTATGATCTTCAATCAATCAACAGAGTTGAAGCT 12315
QY 748 GCAGAGACATTAAGAGAGAAATAGATGATGATCAAGAGAGAGCTGATATATTATTAAC 807
DB 12316 GCTGAGACTATTAAGAGAGAGACAGATGCTGATGATGAAGAGCTGATATATTTCGATC 12375

QY 808 CTAGCTTCTGAACCTATTGCGGCATGTGGGAGAGCTGTATTAACCCATTTGTCAAGAGCT 867
DB 12376 CTTGAGACAGATTTGATTTTCTGCTGTGGAGAACTAGAAAGCTTGAAGTGAAGAGC 12435
QY 868 ATACATGATTAATTAATGAGCATGAGATTTCTTAATTAAGCTTGAAGAGCCGATTTGA 927
DB 12436 ATTATATGATTAATTAATGCTTTGGAGAACTTAACAAACATGAAAGAGAGCTAGA 12494
QY 928 CCAACTTTGAG-----AGCAATGACGCTCCGTCAGTACAGATGAGTGCAGCGGCTATT 981
DB 12495 AAACTTGAGATGCTATGCAAGCTGCTGACAGTATCAGAGACACTCTTCAGGCTATGTT 12554
QY 982 TTTTGGGTGATATTTGACAGTGTGAAGTTGCTTCAATGTCTCAATTTGGAACAGATC 1041
DB 12555 TGACTGCTGATTAACCTGTGATTAACCTTGACCAATGCCCCCTGTGGCACTG-AAC 12613
QY 1042 TCGAACTGTCAACAGACAGATTTGATGTGCTAGAGCAATTTAAGTCTGAGGCTATCAAC 1101
DB 12614 TCAATCTCTTTAAGATCACTTAATTAATGAAGAGGCTTCAAGTGAAGTTTACCAAC 12673
QY 1102 AGCAGATGAGTGTGACACTGACTCATCAAGCAGACGCTTTTGTAAAGAACTAACAG 1161
DB 12674 AGCAAAATTGAGATGAGAGAGCTTAATATCACCAGGGTGAACCTGATTAAGAAAGCTACTG 12733
QY 1162 MAGAGTGAACAAACACACTGTTTACAGCCAAATATGANTGTATATTTGATAT-GGCTA 1220
DB 12734 ATGAGAGGAGCAGAGATTAATTAAGAAACCACTGACAGAACTCAACACTCTGGGAGA 12793
QY 1221 GCCTGAGAGAGAGATCATCAACAGACAGCACTAATCTGAGGGGCTCTATTAAGCCMTGG 1280
DB 12794 ACCGTGGTGAAGAAATTTGCCACAGACAGCAACAACTAAGAGGGCTCTGTGGCCCTTG 12853
QY 1281 GTCAGTTCCAACATGCCCTGTATAGCTCCTGAT--GCTGACACACACCGAGGCTTGC 1338
DB 12854 GTCACTTCCAGCATGCTTATGAGAGCACTAATGATGTGGCTGATCATATCCGAAGATTGT 12913
QY 1339 TAAGTGAACGAAACCTGTTGAGAGACACCTTAAGCCATTTGAATTTGAATTTGCCAACG 1358
DB 12914 TAGATGTCTCAGAGACCAATTAAGTGAAGACCCAAAGTCAATTTGAAGTCTCCAAAGC 12973
QY 1399 ATCATGTCTCCAAATGATGATTAAGCCATCAAGTCAAGTCAAGTGAAGCGTTAATAAG 1458
DB 12974 ACATGTCTTAAATAATGATGATTTTGGCTCATCAAGCACAAGTGAAGAAACATCAACAG 13033
QY 1459 CAGGAAATGATCTTAATGAATCAAGTCAAGAGAGAAAGCAAGCACTTCAGAACAAAGC 1518
DB 13034 CTGGCATGAGCTTCTTGAATCCAGTCTGGAATGATGTCAGAGCACTTAAGAGCCGTT 13093
QY 1519 TAGAGTTTAAATCAACGCTGCAAAATGTTTGGAAAAACAAACAAAGAGAGCAGC 1578
DB 13094 TGGAAAGCATGAACCAATGCTGGAGATCAGTGTTAACGAAAAACAGAGAGGGAGCAGC 13153
QY 1579 AGCTGATGATGCTTGGCGCAGGCAAGGCTTCCATGGGAAATTAAGGATTTGAGAGC 1638
DB 13154 AGCTTCACTCACTCTGAGAGAGCCAGAGGCTTCCAGAGAAATTAAGATTTCTCT 13213
QY 1639 AGTGGTACTGACAGCAGAGAGCTATCTGTTGATCTTAACCGCTGGAGGTTTACCGG 1698
DB 13214 TGGAACTTAAGTAAGAGAGAGCAGCTTCTGATCTTAAGCCACAGAGAGATTTCTG 13273
QY 1699 AAACAGCCAGAGAGAGCTTAATGTCCATATGAGAGTGTGTGCTCTTGAAGCTTAAG 1758
DB 13274 AAACCTGATGAGAAACACTTGTATCAGATATGAAACTCTATTTCCACAGCTGAAGC 13333
QY 1759 AAGAAATTAATTAAGTCTGATGTCAGAAAGGCGACAGATGCTTGAAGATGCCCAAAAT 1818
DB 13334 AAGAGACTTATTAACACTTACTTGAACAGAGGAGACTCAATGCTTGAAGCCGTGACAGCT 13393
QY 1819 CTG 1821
DB 13394 CTG 13396

RESULT 9

ABK34676 ID ABK34676 standard; cDNA; 2181 BP.

XX AC ABK34676;

XX 08-MAY-2002 (first entry)

XX Human cDNA for novel secreted protein, SEQ ID 445.

XX Human; ss: gene; secreted protein; immune deficiency; viral infection;
XX bacterial infection; fungal infection; autoimmune disorder; burn;
XX rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
XX diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
XX Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
XX coagulation disorder; haemophilia; inflammatory disorder; ulcer;
XX tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
XX lymphoid cell deficiency.

XX Homo sapiens.

XX MO200177290-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001MO-US010295.

XX 06-APR-2000; 2000US-0194941P.

XX (GEMV) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
XX Gulkrot K, Graham JR;

XX WPI; 2002-179323/23.

XX Six hundred and twenty five polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for treating
XX immune deficiencies and disorders such as autoimmune disorders.

XX Claim 1; Page 234; 339pp; English.

XX The invention relates to 625 polynucleotides which have been derived from
XX a variety of human tissue sources and sequences that hybridize to them. Also
XX proteins, their complements and sequences that hybridize to them. Also
XX included are a vector comprising the polynucleotide, a host cell
XX transformed with the vector, the proteins encoded by the polynucleotides,
XX antibodies that bind to the proteins and identification of modulators of
XX the proteins or the expression of the polynucleotide. The polynucleotides
XX can be used as probes for the identification and isolation of full length
XX cDNA and genomic DNA. The polynucleotides and proteins can also be used
XX as nutritional supplements. The protein is useful in the treatment of
XX various immune deficiencies and disorders such as viral infections,
XX bacterial infections, fungal infections, autoimmune disorders (e.g.
XX rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and
XX diabetes) and allergic reactions and conditions (e.g. asthma). They are
XX also useful for treating neurodegenerative diseases (e.g. Alzheimer's
XX disease, Parkinson's disease), liver fibrosis, coagulation disorders
XX (e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and
XX tumours. They are also useful for tissue regeneration, for wound healing
XX and in the treatment of burns, incisions and ulcers. The proteins are
XX also useful for regulating haematopoiesis, for treating myeloid or
XX lymphoid cell deficiencies. The present sequence is one of the 625 cDNA
XX sequences encoding a secreted protein

XX Sequence 2181 BP; 679 A; 442 C; 545 G; 515 T; 0 U; 0 Other;

XX Query Match 24.7%; Score 501; DB 6; Length 2181;

XX Best Local Similarity 62.0%; Pred. No. 3,7e-111;

XX Matches 895; Conservative 0; Mismatches 529; Indels 19; Gaps 6;

QY 296 TCATGACAAAGATGATCAGATCCTTGAGAGCCTGGAACGATCGTGAACGCTGAGGC 355

Db 1 TTCATATATAAATTGAGCCTATGTTGAGACACTGAGAAATCTTTCCTCGCTCCGA 60
QY 356 AGCCACCTCTATTCCTGAGAGGTTGAGAAATCAAGAAACGATCAGTAAATTAAG 415
Db 61 TGCCACCACTGATCCTGCTGAAAGTGAAGATCAGAAAGTGCATCACTGACATTAAG 120
QY 416 ATGTGTCACTAGACATGAGAAAGCTACAGCCGTTGTATGAAACTTTAAACAGAGGAG 475
Db 121 GTGCCACCGTGAAGTAAAGAACTGCACCATCTTTGAGGCTTGAAGCCGCTGAG 180
QY 476 AGAAATGATTTGCTAGATCTGGGGGACCTGATTAAGACATCTTCCCAAGCTGTCCAG 535
Db 181 AGAGGTTATTTGACATCTCAGGAGCAGACAAAGATCTGGCTGAAAAGAAATCAAG 240
QY 536 ATAACTTACCAAAATGTTTTCATTTGGAGAAACATACACACACTGTGTGAAGAGAGG 595
Db 241 ATTAATTTGATCAAAATGATTTCTTTGGAGAGACATCAAAAGCTGGGCTGAAGAACAG 300
QY 596 AAGCCAACTACTGATGATGATGAGCTAGACAGAAAAGTTCTGTGTGATCAATGTCAAT 655
Db 301 AATCAAAATTTCTGTATGTCTTGAATTAAGAGAAAGTTCTGTGTATGACATGGCAGCTC 360
QY 656 TGATAGTTACATTAAGATTAATCAAGATTTCAATCCGGACCTGGAAATCTTGAATG 715
Db 361 TCTGACACACATCAAAAGACACCCAGGATATTTGCTATCATCTTGAAGCCAGGATG 420
QY 716 ATCCTTCAGTAAAGAAACAGCAAGAAAGAGAGAGACATTAAGGGAAGAAATGATG 775
Db 421 ATCTTTCATCATCAACAAACAGGTTGAAGCTGTGAGCTTTTAAGAAAGACAGATG 480
QY 776 GACTACAGAGAGAGCTGATATATATTAATTAACCTAGTTCTGAACATCTTGGCGCATG 835
Db 481 GTCTGATGAAAGGCTGAGTTTATTCGATCTTGGACAGATTTGATTTTGGCTGTG 540
QY 836 GGGAGCTGATTAACCATTTGTCAGAAAGATTAATCATATGTTAAATTAAGCATGGAT 895
Db 541 GAGAACTAGAAAGCCTGAAATGAATTAATGCTTGGAGAACTTAAACGAAACATG--- 597
QY 896 TCTCTAAATAAAGCTTGAAGACCGGATGTAACAACTTGAGAGCAATGAGTGCCTG 955
Db 598 ---AAGATAGGCTAGCAAAACCTG-----AGATGCTATGAGGCTGTGTGCG 646
QY 956 AGTACAGATGAGCTGAGCGGATTTTCTGGTATGATTAATTCAGAGTGAATTCGCT 1015
Db 647 TATCAGAGAGCTTCAAGCTATGTTGATCTGCAAGATTAACATGATTAACCTTTC 706
QY 1016 TCAATGTTCCAAATTTGGAACAGATTCGAAACTGTCAAGCAGCATTTGCTTGA 1075
Db 707 ACCAAGCCCCCTGTGGCTCTG-ACCTCGATCTGTTAAAGATCAAGTTAAATGAATGA 765
QY 1076 GCATTTAAGTCTGAGGCTATCAACAGAGATGAGTGTGCACTGATCAACGCA 1135
Db 766 GAGTTTCAAGTAAAGTGTATCCAGAGCGAATTTGATGAGAACTTAAATCAACGAG 825
QY 1136 GAGCTTTTGTCTAAAGAAAGTAAACAGNAGAGTGAACAAACACTGTTCTAGACCCAA 1195
Db 826 TGAAGTATGTTTAAAGAAAGTAAAGTATGAGACGAGACATTAATCGAAGAACACT 885
QY 1196 AATGAGTCTGATTAATGATATGCT-AGCTGAGNAGAGATATCAACAGACAGATTA 1254
Db 886 GACAGAACTCAAAACACTCTGGAGAACCTGGGTGAGAAATTTGCCACCGCAGACAAA 945
QY 1255 ACTGAGGGTGTCTATTTAGCCNTGGGTGAGTTCAACATGCCCCTATAGACTCTCTGA 1314
Db 946 ACTAAGAGGGCTCTGTGGCCCTTGTGTAGTTCAAGATGCTTTAGAGAACTAATGAG 1005
QY 1315 T--GCTGACACACACGAGGCTTGTCTAAGTGAAGAAACCTGTTGAGAGAGACCTTA 1372
Db 1006 TTGGCTGATCTATACGGAAGAGTTGTATGATCTCAGAGAACCAATTAAGTGAAGACCAA 1065
QY 1373 AGCATTTGAATTTGAATCTTGCAGAGCATGTGCTCAAAATGATGATTAAGCCATCA 1432

Db 1066 AGTCATTGAGTTGAGTCGCAGAACACCATGTCTTAAAAATGATGTTTGGCTCATCA 1125
QY 1433 GTCCACAGTGGAGCCGTTAATTAAGCAGAGAAATATATTAATGATCAAGTCAGAGA 1492
Db 1126 AGCCACAGTGGAAAGTCACAAAGCTGGCAATGAGCTTCTTGAATCCAGTCTGAGA 1185
QY 1493 AGAAGCAACCACTTTCAGAACAAAGCTAGAGGTTTAAATCAACCGTCGCAAAATGTTT 1552
Db 1186 TGATGCCACAGCTTAAGAGCCGTTTGAAACCATGAACCAATCTGGAGTCTGTT 1245
QY 1553 GGAAGAAACAGAACAAAGAGAGCAGCAGCTGATGTCCTTGGCCAGCCAAAGGTT 1612
Db 1246 ACAGAAAAACAGAGAGAGAGAGCAGCAGCTTCAATCACTGACAGAGCCAGGCTT 1305
QY 1613 CCATGGCCAAATTGAGATTGTCAGCAGTGGTGACTGACAGAGAGCTCATCTTGGC 1672
Db 1306 CCACAGTAAATTGAAGA-TTCTCTTGAAATTCTGAATGAGAGACCCAGCTTTCGC 1364
QY 1673 ATCTAAACCGCTGGAGGTTTACCGAAACAGCCAGAGAGCAGCTTAATGTCATATGGA 1732
Db 1365 ATCTAAGCCACAGAGAGAGACTTCTGAAACTGTCTGGAACAGCTTGATACATATGCT 1424
QY 1733 AGT 1735
Db 1425 AAT 1427

RESULT 10
AAS29811
ID AAS29811 standard, cDNA, 2113 BP.
AC AAS29811;
XX 21-NOV-2001 (first entry)
XX Human cytoskeletal element-related polypeptide encoding cDNA #42.
DE
XX Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
KW cat; dog; chicken; sheep; immunosuppressive; antiathritic; vasotropic;
KW antirheumatic; antiproliferative; cytoskeletal; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
OS Homo sapiens.
XX
XX WO200155168-A1.
FN
XX
XX 02-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001331.
PE
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0198874P.
PR 17-MAR-2000; 2000US-0199076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI: 2001-476182/51.
XX
XX P-PSDB; AAU18549.
PT
PT Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 1; SEQ ID NO 52; 505bp; English.

CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma. The polypeptides can also be used
CC to aid wound healing, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues and in chemotaxis.
CC The polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at

Query Match 24.5%; Score 497.8; DB 4; Length 2113;
Best Local Similarity 63.7%; Pred. No. 2.2e-110;
Matches 836; Conservative 0; Mismatches 466; Indels 11; Gaps 5;

QY 518 CTGCCAAGCTGTCGAGTAAGCTTGACCAATGGTTTTCATTTGGGAGAAATACAC 577
DB CGGCAAAAGAAATCCGAGTAATATGGATCAATGGTATTTCTTGGAGGACATCAAG 121
QY 578 CACTGTGGAAGAGAGAGAGCCAACTACTGATGTGATGAGCTAGAGAAAAGTTCT 637
DB 122 CTGGGCTGAAGACAGAAATCAATTTCTTGATGTCTCTTAATTAAGAGAAATTTCT 181
QY 638 GGTGTGATCAATGTCATTGATTAATTAACATTAAGATTAAGATTTATCCGGAC 697
DB 182 GGTATGACATGGAGCTCTCTGACACCATCAAAAGACCCAGATATTTGTCATGACT 241
QY 698 TGGAAATCTCGAATTTGATTCCTTCAGTAAATAAAGACGAAGAGACGACACCA 757
DB 242 TGGAAAGCCGAGCATTTGATTCCTTCATCAATCAAAAGATTTGAAGTGCTGAAC 301
QY 758 TAAAGGAAGAAATAGATGAGCTACAGAGAGAGCTGATTAATTAACCTAGTTCTG 817
DB 302 TTAGAGAGAGACAGATGATGCTGCAATGAAGAGCTGGAAGTTATTCGATTCCTTG 361
QY 818 AACTATTCGGGATGTGGAGCCTGATTAACCCATTGTCAAGAGATTAATCATGAGT 877
DB 362 ATTTGATTTTTCCTGTGAGAAACTGAAAGCTGAAGTGAGGAAGACATGATGAG 421
QY 878 TAAATTCAGCAGATGGGATTTCTTAATTAAGCTTGAAGACCGGATTAACCACTTGA 937
DB 422 TGAATATATGCTTGGAGAACTTTAAACAAACATGGAAGAGAGCTTAAGAAACCTTGA 480
QY 938 -----AGCAATGACAGCTCCGTCAGTACAGATGAGACTCAGGCGGTATTTTCTGG 991
DB 481 GATGCTATGCAAGCTCTGTGCAATACAGACACTCTTCAGGCTATGTTGACTGGCT 540
QY 992 GATATTTGACAGTGTAGTTGCTTCATATGTCATTTGAATGGAACAGCATCTCGAA 1051
DB 541 GATTAACATGATTTAACTCTGACACATGCCCCCTGTGGACATG-ACCTCAATCTGT 599
QY 1052 CAAGCAGCAGATTTAGTGTCTAGAGCAATTTAAGTGTGAGGCTTACACAGAGTTAA 1111
DB 600 TAAAGATCAGTTAAATGAATGAATGAAGAGTTCAAGATTAAGAAATTTGA 659
QY 1112 GTGGTACGACTACATCATCAGCAGCGCTTTTGTCTAAAGAAAGTAACAGNAGAGTGA 1171
DB 660 GATGAGAAAGCTTAATACACAGGTTGAATCTGTTTAAAGAAAGCTACTGATGAGCG 719
QY 1172 CAAACACACTGTTTGAACCCAAATTAAGANTGATATTTGATAT-GGTAGCTGGAGNA 1230
DB 720 CAGAGCATTTATTAACGAGAACCACTGACAGAACTCAAAACCTCTGGAGAACCTGGGTGA 779
QY 1231 GAGAAATCAACACAGACAGATTAATGAGAGGTGTCTTAATAGCCNTGGGTCACTTCCA 1290
DB 780 GAAATATGCCCACCGACAGACACAACTAAGAGGGGTCTGTGGCCCTTGTCAGTTTCA 839
QY 1291 AACTGCGCTGTATGAGCTCTGCTGAT--GCTGACACACACGAGGGTGTCTAATGAGCA 1348
DB 840 GATGCTTATGAGAACTAATGATGTTGCTGACTCATACGAAAGATTTTATATCTCTCA 899
QY 1349 GAAACCTGTTGAGAGAGACCTTAAAGCATTTGAATTTGAATCTTCCAAACATCATGTGCT 1408

Db 900 GAGACCAATTAAGTGGAGCACCAAAAGTCATTGAAGTGGAGCTCGCAAAAGCACCATGTCCT 959
Qy 1409 CCAAAATATATGATTAGCCCATCAGTCCACAGTGGAGCCGTTATTAAGCAGAAATGA 1468
Db 960 AAAAAATGATGTTTGGCTCATCAAGCCAGTGAACCAAGCTGGCAATGA 1019
Qy 1469 TCTAATTAATCAAGTGGAGGAAAGAAAGCAACCACTTGAGAACAGTATGAGCTTT 1528
Db 1020 GCTTCTTAATCAGTGTCTGGAGATGATGCCAGCACTTTAGAGACCGTTTGGAGCCAT 1079
Qy 1529 AAATCAACGCTGGCAAAATGTTTGGAAAAACAGAACAAAGAGACAGCACTGATGG 1588
Db 1080 GAACCAATGCTGGAGTCAAGTGTACAGAAAAACAGAGAGAGGAGCAGCACTTCACTC 1139
Qy 1589 TGCCTTGGCCAGGCGCAAAAGGTTCCATGGCGAAATTTAGATTGGCAGCACTGCTAAC 1648
Db 1140 AACTCTGAGAGAGCCAGGCGCTTCACAGTGAATTAAGTTTCTCTTGGAACTTAC 1199
Qy 1649 TGACACGAGGAGCTCATCTGTTGGCATCTAAACCGGTGGAGGTTTACCGAAACAGCCAA 1708
Db 1200 TAGAATGAGAGCCAGCTTTCTGCATCTAAGCCCAAGAGGACTTCTTGAATCTGTAG 1259
Qy 1709 GGAGCAGCTTAATGTCATATGAGAGTGTGCTGCTTTGAGCTAAAGAAACATA 1768
Db 1260 GGAACAGCTTGATACACATATGGAATCTATTCCAGCTGAAGCAAGCAAGAGACTTA 1319
Qy 1769 TAAGAGTGTGATGCAGAAAGCCAGACAGATGCTTGCAAGATGCCAAATCTG 1821
Db 1320 TAATCACTACTTGACAAAGGCGACACTCATGCTTAAAGCCGTGACGACTCTG 1372

RESULT 11
AAS29791
ID AAS29791 standard; cDNA, 5453 BP.
AC AAS29791;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human cytoskeletal element-related polypeptide encoding cDNA #22.
XX
KW Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cyostatic; cardiac; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
XX
OS Homo sapiens.
XX
PN WO20015156-A1.
PD
XX 02-AUG-2001.
PF
XX 17-JAN-2001; 2001WO-US001331.
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235848P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.

20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251909P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476182/51.
DR P-PSDB; AAU18529.
PT
PT Novel isolated human cytoskeletal element-related polypeptide useful for
PT diagnosis/treatment of neoplastic disorders, disorders associated with
PT neural transmission, chromosomal abnormalities, autoimmune disorders.
XX
PS Claim 1; SEQ ID NO 32; 505pp; English.

CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma. The polypeptides can also be used
CC to aid wound healing, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues and in chemotaxis.
CC The polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at

Query Match 24.5%; Score 497.4; DB 4; Length 5453;
Best Local Similarity 63.7%; Pred. No. 4e-110;
Matches 835; Conservative 0; Mismatches 465; Indels 11; Gaps 5;

QY 520 GCCAAGCTGTCAGGATTAAGCTTGACCAATGGTTTATTCATTGGGAGAACATACACACA 579
DB 1 GCAAAAGAAATCCAGATTAATGGATCAATGGTATTTCTTGGAGGACATCAAGCT 60
QY 580 CTGGTGAAGAGAGGAGGAGCAAACTACTGATGTGATGAGCTACAGAAAAGTTCTCG 639
DB 61 CGGGCTGAAGAGAGGAGAAATCAATTTCTGATGTCTTGAATTACAGAGAAATTCTGG 120
QY 640 TGTGATCAGATGTCATTGATTAAGTACATTAAGATCTCAAGATTTCATCCGGACCTG 699
DB 121 TATGACATGGCAGCTCTCTCTGACCAACCAAGAGACCAGGATTTGTCATGACTTG 180
QY 700 GAAGATCTCGAATTAAGTCTTCAATGATTAAGAAACACAGAACACACAGACACATTA 759
DB 181 GAAAGCCAGGCTTATGATCTTTCATCATCAACCAACAGATTGAACTGTGACTTAT 240
QY 760 AGGAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
DB 241 AAGGAAGACACAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 820 CTCATTGCGGATGCGGAGGAGCTGATTAACCATTTGTCAAGAAAGTATACATGATTA 879
DB 301 TTGATTTTTCCTGCTGAGAGAACTGAGAACTGAGAGGAGGATTTGATGAGATG 360
QY 880 AATTACGATGGGATTTCTTAATTAAGCTTGAAGAACCGGATTAACCACTTAG-- 937
DB 361 AATTATGC-TTGGAGAACTTAAACAAACATGAGAAAGAGGCTTGAAGAACTTAGGA 419
QY 938 ---AGCAATGACGCTGCCGTGACATGATGATGATGATGATGATGATGATGATGATGAT 993
DB 420 TGTATGCAAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
QY 994 TATTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1053
DB 480 TAACACTGTGATTAATCTGACCACTGCCCCCTGTGGCACTG-ACCTCAATACGTGTA 538
QY 1054 AGCAGAGATGATGCTAGAGCAATTAAGTGAAGGCTTATCAACGACGATGAGT 1113
DB 539 AAGATCAGTTAATGAATGAAGAGAGTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 598
QY 1114 GGTACAGCTGACTATCAACGACGAGCTTTTCTTAAGAAAGTAAACAGNAGAGATGACA 1173
DB 599 TGGAGAGCTTATATCAACGAGGCTGAAGTATTTAAAGAAACTATGATGAGACGAGACA 658
QY 1174 AACACACTGTTTAAACCAAAATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1232
DB 659 GAGACTTATTAAGAGAACACGACAGCAAGTCAAAACCTCTGGGAGAACTGGGTGAGACA 718
QY 1233 GAATATCAACAGACAGCAATTAAGAGAGGCTCTATTAACGCTGAGTCTGAGTCTCAAC 1292
DB 719 AATATGCCACGACGACGACCAATGAAGAGGAGCTCTGTTGGCCCTTGTGATGATGATGAT 778
QY 1293 ATGCCCTGATAGCTCTGAT--GCTGACACACACGAGGCTTGTCAAGTGAAGACAGA 1350

Db 779 ATGCTTAGAGGAATAATGATGGCTGACTCATACCGAAGAGTTGTTAGATGCTCAGA 838
Qy 1351 AACCTGTTGAGAGAGACCTTAAGCCATTGAATTTAACTTGCACAGCATATGCTCC 1410
Db 839 GACCAATAAGAGGAGACCCAAAGATCATTTGAAGTTGAGCTGCAGAAAGACCATGTCTCAA 898
Qy 1411 AAAATGATGTTATTTAGCCCATGATGCCACAGTGAAGCCGTTAATTAAGACAGAAATATC 1470
Db 899 AAAATGATGTTATTTAGCCCATGATGCCACAGTGAAGCAAGTGAAGCTGCAATGAGC 958
Qy 1471 TAAATGATCAATGAG 1530
Db 959 TTCTTGAATCCAGTGTGAGATGATGCCACAGCTTAAGAGAGCCGTTTGAAGCCATGA 1018
Qy 1531 ATCAACCTGAGCAAAATGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1590
Db 1019 ACCAATCTGGAGATCAGTGTGTTACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
Qy 1591 CCTTGGCCGAGGAG 1650
Db 1079 CTCTGAG 1138
Qy 1651 ACAG 1710
Db 1139 GAATGAG 1198
Qy 1711 AGCAGCTTAATGTCATATGAG 1770
Db 1199 AACAGCTTGATACATATGAG 1258
Qy 1771 AGAGTCTGATGAG 1821
Db 1259 ATCAACTGATCTGAG 1309

RESULT 12

AAL22375 standard; cDNA; 547 BP.

ID AAL22375 standard; cDNA; 547 BP.

XX AAL22375;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 14832.

XX Human; breast cancer; cell marker; cytosstatic; ss.

XX Homo sapiens.

FN WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US000798.

PR 14-JAN-2000; 2000US-0176077P.

PR 14-MAR-2000; 2000US-0189167P.

PR 24-MAR-2000; 2000US-0192099P.

PR 29-MAR-2000; 2000US-0193480P.

PR 15-MAY-2000; 2000US-0205230P.

PR 09-JUN-2000; 2000US-0211315P.

PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.

PS Claim 1; Page 2676; 3695bp; English.

XX

CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX

SQ Sequence 547 BP; 186 A; 103 C; 140 G; 118 T; 0 U; 0 Other;

Query Match 21.6%; Score 438.8; DB 4; Length 547;

Best Local Similarity 94.1%; Pred. No. 2.8e-96;

Matches 508; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

Qy 962 GATGAGCTGAGAGGCGGATTTTCTGGGTAGATATTCAGGTGTAAGTTGCTTCATG 1021

Db 10 GATGAGCTGAGAGGCGGATTTGACTGGGTAGATATTCAGGTGTAAGTTGCTTCATG 69

Qy 1022 TCTCAATTGAGACAGCATCTCGAACTGTCAAGCAGCAGATGATGCTAGAGCAATT 1081

Db 70 TCTCAATTGAGACAG-ATCTCGAACTGTCAAGCAGCAGATGATGAGCAATT 128

Qy 1082 TAACTGAGGCGCTATCAAGCAGATAG-AGTGTAGAGCTGATCAAGCAGAGCGC 1140

Db 129 TAACTGAGGCGCTATCAAGCAGATAGAAATGAAAGAGCTGATCAAGCAGAG-GC 187

Qy 1141 TTTTGCTAAAGAAAGTAAAG 1200

Db 188 TTTTGCTAAAGAAAGTAAAG 247

Qy 1201 AACTGATATTTGAT-AGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259

Db 248 AACTGATATTTGATAGAGATAGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 307

Qy 1260 AGGTGCTCTATTAGCCGCTGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317

Db 308 AGGTGCTCTATTAGCCGCTGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367

Qy 1318 TGACACACACGAGGCGCTGCTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377

Db 368 TGACACACACGAGGCGCTGCTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427

Qy 1378 TTGAATTTGAATTGCAAGCATGCTGCTCAAAATGATTTAGCCCATCAGTCCA 1437

Db 428 TTGAATTTGAATTGCAAGCATGCTGCTCAAAATGATTTAGCCCATCAGTCCA 487

Qy 1438 CAGTGAAGCGGTTAATTAAGCAGAAATGATCTTAATTTGAATCAAGTCAAGAGAGAGAG 1487

Db 488 CAGTGAAGCGGTTAATTAAGCAGAAATGATCTTAATTTGAATCAAGTCAAGAGAGAGAG 547

RESULT 13

AAL13506 standard; cDNA; 531 BP.

ID AAL13506 standard; cDNA; 531 BP.

AC AAL13506;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 5963.

XX Human; breast cancer; cell marker; cytosstatic; ss.

XX Homo sapiens.

FN WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US000798.

XX

PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX
XX Claim 1; Page 1070-1071; 3695pp; English.
XX
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AA07544-AA026789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity
XX
XX Sequence 531 BP; 179 A; 102 C; 133 G; 117 T; 0 U; 0 Other;

Query Match 20.8%; Score 422.8; DB 4; Length 531;
Best Local Similarity 93.9%; Pred. No. 2.1e-92;
Matches 492; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

QY 962 GATGACCTGAGGCGGATTTTCTGGGTAGATTTGCAAGTGTGAATGCTTCAATG 1021
DB 10 GATGAGCTGAGGCGGATTTTCTGGGTAGATTTGCAAGTGTGAATGCTTCAATG 69
QY 1022 TCTCCAAATGGAACAGATCTCGAAACTGTCAAGCAGACAGATGTGTAGAGCAAT 1081
DB 70 TCTCCAAATGGAACAG-ATCTCGAAACTGTCAAGCAGACAGATGTGTAGAGCAAT 128
QY 1082 TAACTGAGGCTTATCAACAGAGATAG-AGTGTAGCACTGACTCATCAGCAGACGC 1140
DB 129 TAACTGAGGCTTATCAACAGAGATAGAAATGGAAGCTGAATCATCAGACAG-GC 187
QY 1141 TTTTGTAAAGAAAGTAAACAGAGAGTAAACACACTGTTCTAGACCCAAATATG 1200
DB 188 TTTTGTAAAGAAAGTAAACAGAGAGTAAACACACTGTTCTAGACCCCAATATG 247
QY 1201 ANCTGATATGATAT-GGCTAGGCTGAGAGAGAAATCATCAACAGACAGATTAATG 1259
DB 248 AACTGAATGATATGATAGGAGTGGCTGAGAGAGAAATCATCAACAGACAGATTAATG 307
QY 1260 AGGCTCTATTATAGGCTGAGGCTGAGTTCACATGCGCTGTATGAGCTCTG--CATG 1317
DB 308 AGGCTCTATTATAGGCTGAGGCTGAGTTCACATGCGCTGTATGAGCTCTGAGTGG 367
QY 1318 TGAACACACCGAGGCTGCTTAAGTGAAGAGAACTGTGAGAGAGACCTTAAGCCCA 1377
DB 368 TGAACACACCGAGGCTGCTTAAGTGAAGAGAACTGTGAGAGAGACCTTAAGCCCA 427
QY 1378 TTGAATTTGAATCTGCAAGCATCATGTGCTCCAAATGATGTATTAGCCCATCAGTCCA 1437
DB 428 TTGAATTTGAATCTGCAAGCATCATGTGCTCCAAATGATGTATTAGCCCATCAGTCCA 487
QY 1438 CAGTGAAGCCGTTAATAAGAGAGAAATGATTAATTGAATCA 1481
DB 488 CAGTGAAGCCGTTAATAAGAGAGAAATGATTAATTGAATCA 531

RESULT 14
AA65919

ID AAF65919 standard; cDNA; 401 BP.
XX
XX AAF65919;
AC
XX 09-APR-2001 (first entry)
DT
XX
XX Novel human polynucleotide, SEQ ID NO: 1675.
DE
XX
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; 88.
XX
XX Homo sapiens.
XX
XX MO200102568-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000MO-US018374.
XX
XX 02-JUL-1999; 99US-0142310P.
XX
XX 02-JUL-1999; 99US-0142311P.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LF, Escobedo J, Imitis MA, Garcia PD, Klinger J;
XX Kassar A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
XX Drmanac R, Ckenjakov R, Drmanac S, Dickson M, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI; 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
XX mammalian cell and detecting cancer, particularly of the colon or
XX prostate, comprises 3351 human polynucleotide sequences.
XX
XX Claim 9; Page 784; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
XX polynucleotides. The library is used to detect differentially expressed
XX genes correlated with a cancerous state of a mammalian cell and can
XX detect colon, prostate, breast and lung cancer. The library can be used
XX to produce probes for detection of mRNA and to produce additional copies
XX of the polynucleotides. The probes can be used for chromosome mapping of
XX the polynucleotide and for detection of transcription levels. Ribozymes
XX or antisense oligonucleotides can be generated. The polynucleotides and
XX their gene products are used as genetic or biochemical markers (e.g. in
XX blood or tissues) that will detect the earliest changes along the
XX carcinogenesis pathway and/or monitor the efficacy of therapies and
XX CC preventive interventions. The polynucleotides, polypeptides and
XX CC antibodies against them can be used in pharmaceutical compositions to
XX treat the cancers and proliferative disorders such as neoplasia,
XX dysplasia and hyperplasia
XX

Seq Sequence 401 BP; 128 A; 75 C; 112 G; 85 T; 0 U; 1 Other;
Query Match 17.6%; Score 357; DB 5; Length 401;
Best Local Similarity 97.3%; Pred. No. 1.8e-76;
Matches 363; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 291 TCAGTTCATGACAAGATGATCAATCTTGTAGAGCCCTGGAACGATCGTGAACGTCT 350
DB 28 TCAGTTCATGACAAGATGATCAATCTTGTAGAGCCCTGGAACGATCGTGAACGTCT 87
QY 351 GAGGACGCCACCTCTATCTTGCAGAGGTTGAGAAATCAAGAAACAGATGAGTGAANA 410
DB 88 GAGGACGCCACCTCTATCTTGCAGAGGTCGTGAAGATCAAGAAACAGATGAGTGAANA 147
QY 411 TAAAGATGTGTAGATGAGAGATGAGAAAGCTAAGCCGTTGTGAACCTTTAAACAGAG 470
DB 148 TAAAGATGTGTAGATGAGAGATGAGAAAGCTAAGCCGTTGTGAACCTTTAAACAGAG 207
QY 471 GGAAGAGAAATGATTTGATCTGGGGGAGACTGATAAAGACATATCTGCCAAAGCTGT 530

Db 208 GGGAGAGAAATGATTGCTAGATCTGGGGGAGCTGATTAACATATGTGCAAGCTGC 267
 Qy 531 TCAGATTAAGCTTGAACCAATGTTTTCATTTGGAGAACATACACACTGTGGAGA 590
 Db 268 TCAGATTAAGCTTGAACCAATGTTTTCATTTGGAGAACATACACACTGTGGAGA 327
 Qy 591 GAGGGAAGCCAACTACTGATGTGATGAGCTAGCCGAAAGTTTGTGTGATTCAT 650
 Db 328 GAGGGAAGCCAACTACTGATGTGATGAGCTAGCCGAAAGTTTGTGTGATTCAT 387
 Qy 651 GTCAATTGATGTT 663
 Db 388 GTCAATTGATGTT 400
 RESULT 15
 AAK51484
 ID AAK51484 standard; cDNA; 4327 BP.
 AC AAK51484;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 29.
 XX
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00653561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Aeundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR P-PSDB; AAM78351.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 1; Page 570-575; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 4327 BP; 1365 A; 921 C; 1046 G; 995 T; 0 U; 0 Other;
 Query Match 17.3%; Score 352.2; DB 4; Length 4327;
 Best Local Similarity .94.9%; Pred. No. 6.5e-75;
 Matches 374; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
 Qy 1494 GAGCAAGCAACCTTCAGAACAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTG 1553
 Db 1 GCACGAGGCAACCTTCAGAACAGCTAGAGTTTAAATCAACGCTGGCAAAATGTTTG 60
 Qy 1554 GAAAAACAGAAACAAAGAGAGCAAGCTAGAGTGTGCTGGCCGAGCAAAAGGTTTC 1613
 Db 61 GAAAAACAGAAACAAAGAGAGCAAGCTAGAGTGTGCTGGCCGAGCAAAAGGTTTC 119
 Qy 1614 CATGGCGAAATTGAGATTGAGATTGAGCAAGCTGCTGACTGACACGAGCGTCAATCTGTTGGCA 1673
 Db 120 CATGGCGAAATTGAGATTGAGATTGAGCAAGCTGCTGACTGACACGAGCGTCAATCTGTTGGCA 179
 Qy 1674 TCTAACCGCTGGAGGTTTACCGGAAACAGCCAGAGAGCAAGCTTAATGTCATATGGA 1733
 Db 180 TCTAACCGCTGGAGGTTTACCGGAAACAGCCAGAGAGCAAGCTTAATGTCATATGGA 239
 Qy 1734 GTCTGTGCTGCTTGAAGCTTAAGAGAAACATATAAGTCTGATGCAAAAGGCCAG 1793
 Db 240 GTCTGTGCTGCTTGAAGCTTAAGAGAAACATATAAGTCTGATGCAAAAGGCCAG 299
 Qy 1794 CAGATGCTTGCAAGATGCCCAAAATCTGACAGACAAATATTGACCAAGACATTAATTAAC 1853
 Db 300 CAGATGCTTGCAAGATGCCCAAAATCTGACAGACAAATATTGACCAAGACATTAATTAAC 359
 Qy 1854 TTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1887
 Db 360 TTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 393

Search completed: January 12, 2005, 08:50:38
 Job time : 932 secs

inis Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: January 12, 2005, 11:10:16 ; Search time 1086 Seconds
(without alignments)
10745.764 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031
Sequence: 1 ctgcacatgcgcagcagaaca.....tngagagaccacacgcgc 2031

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2024	99.7	2031	15	US-10-089-887-46
2	1656.8	81.6	9274	9	US-09-885-535-3
3	1653.6	81.4	16684	18	US-10-723-860-7690
4	1637.6	80.6	3307	16	US-10-108-260A-1453
5	807	39.7	3987	14	US-10-198-846-12468
6	652.2	32.1	18529	14	US-10-198-846-12599
7	501	24.7	2181	9	US-09-823-245A-445
8	497.8	24.5	2113	15	US-10-158-034-52
9	497.4	24.5	5453	15	US-10-158-034-32
10	430.8	21.2	599	14	US-10-198-846-4738
11	266.8	13.1	4495	15	US-10-037-270-352
12	266.8	13.1	4495	15	US-10-117-722-352

13	262.8	12.9	344	11	US-09-864-408A-111	Sequence 111, App
14	235.6	11.6	1721	14	US-10-043-487-126	Sequence 126, App
15	177.8	8.8	3684	15	US-10-094-749-272	Sequence 272, App
16	134.6	6.6	563	15	US-10-029-386-4198	Sequence 4198, App
17	133.6	6.6	171	15	US-10-029-386-17898	Sequence 17898, App
18	93.8	4.6	1024	13	US-10-202-193-9	Sequence 9, App1
19	88.6	4.4	449	15	US-10-158-034-51	Sequence 51, App1
20	88.6	4.4	4156	15	US-10-158-034-31	Sequence 31, App1
21	65.2	3.2	3720	10	US-09-822-846-180	Sequence 180, App
22	63.4	3.2	845	14	US-10-043-487-125	Sequence 125, App
23	63.4	3.1	364	14	US-10-198-846-9302	Sequence 9302, App
24	63.2	3.1	2117	15	US-10-104-047-16	Sequence 16, App1
25	61.2	3.0	361	14	US-10-198-846-8468	Sequence 8468, App
26	61	3.0	1595	15	US-10-369-493-28935	Sequence 28935, App
27	61	3.0	1595	15	US-10-369-493-28937	Sequence 28937, App
28	60	3.0	60	10	US-09-808-975-12022	Sequence 12022, App
29	58.8	2.9	3510	10	US-09-845-416-12	Sequence 12, App1
30	58.8	2.9	4476	10	US-09-845-416-31	Sequence 31, App1
31	58.6	2.9	663	18	US-10-425-115-86406	Sequence 86406, App
32	58.2	2.9	923	16	US-10-343-710-16	Sequence 16, App1
33	57.6	2.8	450	15	US-10-312-495-4	Sequence 4, App1
34	57.6	2.8	2048	10	US-09-814-353-20207	Sequence 20207, App
35	56.8	2.8	361	14	US-10-198-846-8619	Sequence 8619, App
36	56.2	2.8	14568	15	US-10-311-455-203	Sequence 203, App
37	55.2	2.7	139	18	US-10-425-115-80204	Sequence 80204, App
38	55.2	2.7	716	9	US-09-772-134B-81	Sequence 81, App1
39	54	2.7	2357	15	US-10-106-698-2037	Sequence 2037, App
40	53.4	2.6	1352	13	US-10-080-960-7	Sequence 7, App1
41	53	2.6	1947	15	US-10-171-581-135	Sequence 135, App
42	52.8	2.6	674	16	US-10-168-948B-5	Sequence 5, App1
43	52.8	2.6	1629	18	US-10-425-115-97514	Sequence 97514, App
44	52.8	2.6	2519	13	US-10-139-262-9	Sequence 9, App1
45	52.8	2.6	2519	14	US-10-253-969-9	Sequence 9, App1

ALIGNMENTS

RESULT 1

US-10-089-887-46

Sequence 46, Application US/10089887

Publication No. US20030219740A1

GENERAL INFORMATION:

APPLICANT: Bayer Corporation et al.

TITLE OF INVENTION: DNA Sequences Isolated from Human Colonic Epithelial Cells

FILE REFERENCE: 1657/1020

CURRENT APPLICATION NUMBER: US/10/089, 887

CURRENT FILING DATE: 2000-08-08

PRIOR APPLICATION NUMBER: US 60/147, 933

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 61

SOFTWARE: PatentIn version 3.1

SEQ ID NO 46

LENGTH: 2031

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (1162)..(1162)

OTHER INFORMATION: Unknown

FEATURE:

NAME/KEY: misc feature

LOCATION: (1202)..(1202)

OTHER INFORMATION: Unknown

FEATURE:

NAME/KEY: misc feature

LOCATION: (1225)..(1229)

OTHER INFORMATION: Unknown

FEATURE:

NAME/KEY: misc feature

LOCATION: (1277)..(1277)

OTHER INFORMATION: Unknown

FEATURE:

/ NAME/KEY: misc feature
/ LOCATION: (1885) .. (1885)
/ OTHER INFORMATION: Unknown
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1889) .. (1889)
/ OTHER INFORMATION: Unknown
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (2013) .. (2013)
/ OTHER INFORMATION: Unknown
US-10-089-887-46

Query Match 99.7%; Score 2024; DB 15; Length 2031;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2031; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCCATGCTGACAGAAACCAATCAATCTCTCAGCTCCGCCGCCAGCCCTTGAAATA 60
DB 1 TTGCCATGCTGACAGAAACCAATCAATCTCTCAGCTCCGCCGCCAGCCCTTGAAATA 60
QY 61 TGAATCTTAAGGACAGAGAGAGAAACATTCGGCAACTCGGAGTTGATTAAGTAAACA 120
DB 61 TGAATCTTAAGGACAGAGAGAGAAACATTCGGCAACTCGGAGTTGATTAAGTAAACA 120
QY 121 CAAGCTCATATAGATTAAGTAAACAACTGGCCACAGTACTGAATGAAGCCCTGGGGAA 180
DB 121 CAAGCTCATATAGATTAAGTAAACAACTGGCCACAGTACTGAATGAAGCCCTGGGGAA 180
QY 181 GGGCTTTCTATCCAGAGAAAGTATGTGGCAGCCGACACCTTTTACAGTCAATTAAAGAA 240
DB 181 GGGCTTTCTATCCAGAGAAAGTATGTGGCAGCCGACACCTTTTACAGTCAATTAAAGAA 240
QY 241 GATGTCAAAAAGGCTGTGTGGCACTGGATGAAGCCATTCTCAATCACTCACTTCCT 300
DB 241 GATGTCAAAAAGGCTGTGTGGCACTGGATGAAGCCATTCTCAATCACTCACTTCCT 300
QY 301 GACAAATAGATCAGATCCTTGAAGAGCTTGAAGAGCATCGTGAACGCTGAGGACGCA 360
DB 301 GACAAATAGATCAGATCCTTGAAGAGCTTGAAGAGCATCGTGAACGCTGAGGACGCA 360
QY 361 CCTCTATCTCTCAGAGAGTTGAGAAAGATCAAGGAAACAGATCACTGAGAAATTAAGATGTG 420
DB 361 CCTCTATCTCTCAGAGAGTTGAGAAAGATCAAGGAAACAGATCACTGAGAAATTAAGATGTG 420
QY 421 TCAGTGAACAATGAGAAAGCTACAGCGCTGTATGAACCTTTTAAACAGAGGGAGAGGAA 480
DB 421 TCAGTGAACAATGAGAAAGCTACAGCGCTGTATGAACCTTTTAAACAGAGGGAGAGGAA 480
QY 481 ATGATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAAG 540
DB 481 ATGATTGCTAGATCTGGGGGGAAGCTGATTAAGACATATCTGCCAAAGCTGTTCAAGATAAG 540
QY 541 CTGGACCAAAATGGTTTTTCAATTTGGAGAACATACACACTGGTGAAGAGAGGAAAGCC 600
DB 541 CTGGACCAAAATGGTTTTTCAATTTGGAGAACATACACACTGGTGAAGAGAGGAAAGCC 600
QY 601 AAACCTGATGATGATGAGAGCTAGAGAAAGTTTCTGATGATCAATGATTAAGTAAAG 660
DB 601 AAACCTGATGATGATGAGAGCTAGAGAAAGTTTCTGATGATCAATGATTAAGTAAAG 660
QY 661 GTTACATTTAAAGATCACTCAAGATTTCACTCGGAGACTGGAGATCTTGGAATTTGATCCT 720
DB 661 GTTACATTTAAAGATCACTCAAGATTTCACTCGGAGACTGGAGATCTTGGAATTTGATCCT 720
QY 721 TCAGTATGTAACCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 TCAGTATGTAACCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CAGGAGAGAGCTGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 840
DB 781 CAGGAGAGAGCTGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 840

QY 841 CCTGATAAACCAATTTCTCAAGAGAGATATACAGTTAAATTCAGCGATGGGATTTCTCT 900
DB 841 CCTGATAAACCAATTTCTCAAGAGAGATATACAGTTAAATTCAGCGATGGGATTTCTCT 900
QY 901 AAATTAAGCTTGGAAAGAACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTCAGTAC 960
DB 901 AAATTAAGCTTGGAAAGAACCGGATTTGACCAACTTGAAGCAATGACGCTGCCGTCAGTAC 960
QY 961 AGATGAGCTGACAGGCGGTTATTTTCTGGGTAGATATTTGAGAGGATGATGAGTGGCTTCAAT 1020
DB 961 AGATGAGCTGACAGGCGGTTATTTTCTGGGTAGATATTTGAGAGGATGATGAGTGGCTTCAAT 1020
QY 1021 GTCTCCAAATTTGAAACAGCATCTCGAAACTGTCAAGAGAGAGATTTGATGATGACCAAT 1080
DB 1021 GTCTCCAAATTTGAAACAGCATCTCGAAACTGTCAAGAGAGAGATTTGATGATGACCAAT 1080
QY 1081 TTAAGTCTGAGGCTTATCAACAGAGATAGAGTGTGACAGTCACTCATCAGCAGAGCCG 1140
DB 1081 TTAAGTCTGAGGCTTATCAACAGAGATAGAGTGTGACAGTCACTCATCAGCAGAGCCG 1140
QY 1141 TTTTGTAAAGAAAGTAAAGNAGAGAGTGAACAACACTGTTCTAGACCCCAAAATAGG 1200
DB 1141 TTTTGTAAAGAAAGTAAAGNAGAGAGTGAACAACACTGTTCTAGACCCCAAAATAGG 1200
QY 1201 ANCTGATATTGATATAGCTAGCTGAGAGAGAGATCATCAACAGACATTAATCTGGA 1260
DB 1201 ANCTGATATTGATATAGCTAGCTGAGAGAGAGATCATCAACAGACATTAATCTGGA 1260
QY 1261 GGGTGTCTTATTAAGCCNTGGGTCAAGTCCCAACATGCTCTGTATGAGCTCTGATGCTGA 1320
DB 1261 GGGTGTCTTATTAAGCCNTGGGTCAAGTCCCAACATGCTCTGTATGAGCTCTGATGCTGA 1320
QY 1321 CACACACGAGGGCTTGTCAATGAGAGAGAACTGTGAGAGAGAGCCCTAAAGGCAATTG 1380
DB 1321 CACACACGAGGGCTTGTCAATGAGAGAGAACTGTGAGAGAGAGCCCTAAAGGCAATTG 1380
QY 1381 AAATTAAGCTTCCAAAGCATCATGTGCTCCAAATGATATTAAGCCATCACTGATCCAG 1440
DB 1381 AAATTAAGCTTCCAAAGCATCATGTGCTCCAAATGATATTAAGCCATCACTGATCCAG 1440
QY 1441 TGGAGCCGTTAATTAAGAGAGAAATGATCTAATTAATCAAGTCAAGAGAGAGAGAGAG 1500
DB 1441 TGGAGCCGTTAATTAAGAGAGAAATGATCTAATTAATCAAGTCAAGAGAGAGAGAGAG 1500
QY 1501 GCAACCTTGAAGAAAGCTAGAGAGTTTAAATCAACGCTGGGAGAGAGAGAGAGAGAG 1560
DB 1501 GCAACCTTGAAGAAAGCTAGAGAGTTTAAATCAACGCTGGGAGAGAGAGAGAGAGAG 1560
QY 1561 CAGAACAAAG 1620
DB 1561 CAGAACAAAG 1620
QY 1621 AAATTAAGAGATTTGACAGAGTGGCTGATCAACGAGAGCTGATCTGTTGGCATTTAAAC 1680
DB 1621 AAATTAAGAGATTTGACAGAGTGGCTGATCAACGAGAGCTGATCTGTTGGCATTTAAAC 1680
QY 1681 CGCTGGAGAGTTTACCGGAAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1740
DB 1681 CGCTGGAGAGTTTACCGGAAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1740
QY 1741 CTGCTTTGAAAGCTAAAG 1800
DB 1741 CTGCTTTGAAAGCTAAAG 1800
QY 1801 TTTGCAAGATGCCCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB 1801 TTTGCAAGATGCCCAAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 AA 1920
DB 1861 AA 1920
QY 1921 AATTTAGACCTGCGCGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980

Db	1921	AATCTTAACCTGCGCCGGCGCGCGCTTGAGCCTTATGTAGTCATTATGAATGAA	1980
Oy	1981	TCACTAGTCGCGCCGCCCTGCAGGTCAACCATATNGAGAGCTCCCAACCGCT	2031
Db	1981	TCACTAGTCGCGCCGCCCTGCAGGTCAACCATATNGAGAGCTCCCAACCGCT	2031
 RESULT 2 US-09-885-535-3			
; Sequence 3, Application US/09885535			
; Patent No. US20020104105A1			
; GENERAL INFORMATION:			
; APPLICANT: Myriad Genetics, Inc.			
; APPLICANT: Helichman, Karen			
; APPLICANT: Bartel, Paul L.			
; TITLE OF INVENTION: Protein-Protein Interactions			
; FILE REFERENCE: 2318-266-II			
; CURRENT APPLICATION NUMBER: US/09/885,535			
; CURRENT FILING DATE: 2001-06-21			
; PRIOR APPLICATION NUMBER: US 60/213,245			
; PRIORITY FILING DATE: 2000-06-22			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: Patentin version 3.0			
; SEQ ID NO 3			
; LENGTH: 9274			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-885-535-3			
 Query Match 81.6%; Score 1656.8; DB 9; Length 9274; Best Local Similarity 95.3%; Pred. No. 0; Matches 1805; Conservative 0; Mismatches 72; Indels 18; Gaps 9;			
Oy	8	GCTGACAGAAACAATATCATATCTCTCACG-TCCGCCCCAGCCCTTGAAATATGATAC	66
Db	3612	GCTGACAGAAACAATATCATATCTCTCACGTTCCGCCCGACCCCTTGAAATATGAAC	3671
Oy	67	TCTAAGGACGACGACGAGAAGAACATCGGCAACTGCGTAGTTGATAGCTGAACACAGCC	126
Db	3672	TCTAAGGACGACGACGAGAAGAACATCGGCAACTGCGTAGTTGATAGCTGAACACAGCC	3731
Oy	127	TCATATAGATAAGATGAACAAAACCTGG---CCACAGTACGAAATGAGCCCTGGGGAAG	182
Db	3722	TCATATAGATAAGATGAACAAAACCTGGGCGACAGTTACTGGAATTGAACCCCTGGGGAAG	3791
Oy	183	CTTTTCTATCCAAGAGAGTATGTGGCAGCGCAKACCCTTTNCACTCATTAAGAAGA	242
Db	3792	CTTTTCTATCCAAGAGAGTATGTGGCAGCGCAKACCCTTTNCACTCAAATTAAGAAGA	3851
Oy	243	TGTCAAAAAGCGTGTGTGGGCACTGGATGAAGCCATTTCTCAATCAACTCAAGTTCATGA	302
Db	3852	TGTCAAAAAGCGTGTGTGGGCACTGGATGAAGCCATTTCTCAATCAACTCAAGTTCATGA	3911
Oy	303	CAAGATAGTCAGATCTTTGAGAGCGTGGAAAGCATCGTGNAACTCTGAGGCAAGCC	362
Db	3912	CAAGATAGTCAGATCTTTGAGAGCGTGGAAAGCATCGTGNAACTCTGAGGCAAGCC	3971
Oy	363	CTCTATCTCTGACAGGTTGAGAGATCAAGAACAGATCAAGTCAAGTAAATTAAGATGTGC	422
Db	3972	CTCTATCTCTGACAGGTTGAGAGATCAAGAACAGATCAAGTCAAGTAAATTAAGATGTGC	4031
Oy	423	AGTAAACATGGAAGAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGGAAT	482
Db	4032	AGTAAACATGGAAGAGCTACAGCCGTTGTATGAACCTTTAAACAGAGGGGAGAGGAAT	4091
Oy	483	GATTCTAGATTTGGGGGAGCTGATPAAAGACATATCTGCGAAAGCTGTTCAAGATACCT	542
Db	4092	GATTCTAGATTTGGGGGAGCTGATPAAAGACATATCTGCGAAAGCTGTTCAAGATACCT	4151
Oy	543	TGACCAATATGTTTTATTATTTGGGAGAACATACACACACTGTGTGAAGAGAGGAACCA	602
Db	4152	TGACCAATATGTTTTATTATTTGGGAGAACATACACACACTGTGTGAAGAGAGGAACCA	4211

QY	603	ACTACTGAGTGTGATGGAAGCTAGCAGAAAAGTTCTGGTGTGATCACTGTCTATTGATAGT	662
Db	4212	ACTACTGAGTGTGATGGAAGCTAGCAGAAAAGTTCTGGTGTGATCACTGTCTATTGATAGT	4271
QY	663	TACCAATTAAATATCTCAAGATTTTCATCCGGGACCTGGAAAGATCTCTGGAAATGTATCCCTTC	722
Db	4272	TACCAATTAAATATCTCAAGATTTTCATCCGGGACCTGGAAAGATCTCTGGAAATGTATCCCTTC	4331
QY	723	AGTAGTAAAAAACAAGCAGAAAGCAGAGAGACCATTAAAGGAGAGAAAATAGATGCACTACA	782
Db	4332	AGTAGTAAAAAACAAGCAGAAAGCAGAGAGACCATTAAAGGAGAGAAAATAGATGCACTACA	4391
QY	783	GGAGAGCTGTGATATAGTTATTAACTTAGGTTCTGAACTCATTTGGCGCATGTGGGAGCC	842
Db	4392	GGAGAGCTGTGATATAGTTATTAACTTAGGTTCTGAACTCATTTGGCGCATGTGGGAGCC	4451
QY	843	TGATTAACCCCTTGTCAAGAAAGACTATACATGACTTTAAATTCAAGCGATGGGATTCCTTAA	902
Db	4452	TGATTAACCCCTTGTCAAGAAAGACTATAGATTAATTAATTCACG - ATGGGATTCCTTAA	4510
QY	903	ATTAAGCTTGGAAAAGACCGGATTTACCAACTTGAAG-----AGCAATGCAGCTGCCGTCA	956
Db	4511	ATTAAGCTTGGAAAAGACCGGATTTACCAACTTGAAGAGCGCAATCAGGCTGCCCTTCACT	4570
QY	957	GTAACAGATGACTGCAAGCGGTAATTTTTCTGGGTAGATATTGACAGTGTGTAAGTTTCGTT	1016
Db	4571	ACCAGGATGACTGCAAGCGGTAATTTTTCTGGGTAGATATTGACAGTGTGTAATTTAGCTT	4630
QY	1017	CAATGTCTCCCAATTGGAAACAGACTCTGAAACTGTCAAGCAGCAGATTTGATGTGTCTAGAG	1076
Db	4631	CAATGTCTCCCAATTGGAAACAG - ATCTCGMAACTGTCAAGCAGCAGATTTGAAGACTTAAG	4689
QY	1077	CAATTTAAGCTGAGGCCCTATCAACAGCAGATAG - AGTGTATGACATGCACTCATCAGCA	1135
Db	4680	CAATTTAAGCTGAGGCCCTATCAACAGCAGATAGAAATGAAAATCTGATATCATCAAGCA	4749
QY	1136	GACGCTTTTGTGTAAGAAAGTATACAGNAGAGATGACAAACACACTGTCTTAGACCCAAA	1195
Db	4750	GA - GCTTTTGTGTAAGAAAGTATACAGAAAGAGTGAACAAACACACTGTCTTAGACCCATT	4808
QY	1196	AATGAGATGTATTTGATAT - GGCTAGCTGTGAAGNAGAAATCATCAACAGACAGCATTA	1254
Db	4809	AATGAGATGTAAATTTGATATGAGATAGCTGTGAAGAGAGAAATCATCAACAGACAGCATTA	4868
QY	1255	ACTGAGGGGTCTCTAATTAGCCGNTGGGGTCACTTCCAAACATGCCCTGTATGAGTCCGAG - -	1312
Db	4869	ACTGAGGGGTCTCTAATTAGCCCTTGGGGTCACTTCCAAACATGCCCTGTATGAGTCCGAGT	4928
QY	1313	CATGCTGACACACACCGAGGGCTTGTCTAATGAGACAGAAACCTTGTGAGAGAGACCCTTAA	1372
Db	4929	ATGCTGTGACACACACCGAGGGCTTGTCTAATGAGACAGAAACCTTGTGAGAGAGACCCTTAA	4988
QY	1373	AGCCATTGAATTGAATCTTGCCAGACATCATGTCTCTCCAAAATATGATATTAGCCCATCA	1432
Db	4989	AGCCATTGAATTGAATCTTGCCAGACATCATGTCTCTCCAAAATATGATATTAGCCCATCA	5048
QY	1433	GTCCCACTGTGAAGCCGTTAATTAAGCAGGAAATGATCTTAATTAATCAACTGACAGAGA	1492
Db	5049	GTCCCACTGTGAAGCCGTTAATTAAGCAGGAAATGATCTTAATTAATCAACTGACAGAGA	5108
QY	1493	AGAACAGACAACTTCAAGACCAAGCTTGAAGTTTAAATCAACGCTGGCAAAATGTTTT	1552
Db	5109	AGAACAGACAACTTCAAGACCAAGCTTGAAGTTTAAATCAACGCTGGCAAAATGTTTT	5168
QY	1553	GGAAAAAACAAGAACAAAGAAAGCAGCAGCTGAATGTGTCTTGGCCCAAGCCAAAGGTT	1612
Db	5169	GGAAAAAACAAGAACAAAGAAAGCAGCAGCTGAATGTGTCTTGGCCCAAGCCAAAGGTT	5228
QY	1613	CCATGGCGAAATTGAGATTTTGCAGCAGTGGCTGTACGACCGAGAGCTCATCTGTGGCC	1672
Db	5229	CCATGGCGAAATTGAGATTTTGCAGCAGTGGCTGTACGACCGAGAGCTCATCTGTGTGGC	5288

QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACGCCAAGACAGCTTAATGTCATATGGA 1732
| | | | |
Db 5289 ATCTAAACCGCTGGAGGTTTACCGGAAACGCCAAGACAGCTTAATGTCATATGGA 5348
| | | | |
QY 1733 AGTCTGCTGCTCTTTTAACTAAAGAAACATTAAGTCTGATGCAAGAAAGGCCA 1792
| | | | |
Db 5349 AGTCTGCTGCTCTTTTAACTAAAGAAACATTAAGTCTGATGCAAGAAAGGCCA 5408
| | | | |
QY 1793 GCAGATGCTTGAAGATGCCCCAAATCTGAGAGACAAATATTGACCAAGACATTAATTA 1852
| | | | |
Db 5409 GCAGATGCTTGAAGATGCCCCAAATCTGAGAGACAAATATTGACCAAGACATTAATTA 5468
| | | | |
QY 1853 CTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAC 1887
| | | | |
Db 5469 CTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAC 5503
| | | | |
RESULT 3
US-10-723-860-7690
; Sequence 7690, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7690
; LENGTH: 16684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (857)..(877)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1454)..(1488)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1960)..(2049)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3121)..(3143)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4680)..(4703)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6993)..(7011)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10458)..(10533)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13384)..(13422)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13686)..(13756)

; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15350)..(15368)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7690

Query Match 81.4%; Score 1653.6; DB 18; Length 16684;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 74; Indels 18; Gaps 9;

QY 8 GCTGACAGAAACCAATCAATCATCTCTGAGC-TCCGCGCCAGCCCTTGAATATGATAC 66
| | | | |
Db 10732 GCTGACAGAAACCAATCAATCATCTCTGAGCTCCGCGCCAGCCCTTGAATATGAAAC 10791
| | | | |
QY 67 TCTAAGGACGACGACGAGAAACATCTGGCACTGCTGTGATGATAGCTGAACACAAAGC 126
| | | | |
Db 10792 TCTAAGGACGACGACGAGAAACATCTGGCACTGCTGTGATGATAGCTGAACACAAAGC 10851
| | | | |
QY 127 TCATATAGATTAAGATGAACAAACTGG---CCACAGTACTGAATGAGCCCTGGGGAAG 182
| | | | |
Db 10852 TCATATAGATTAAGATGAACAAACTGGCCACAGTTACTGGAATTGAGCCCTGGGGAAG 10911
| | | | |
QY 183 CTTTCTATCCAAGAGATATGTGACGACCGACACCCCTTACAGTCACATTAAAGAGA 242
| | | | |
Db 10912 CTTTCTATCCAAGAGATATGTGACGACCGACACCCCTTACAGTCACATTAAAGAGA 10971
| | | | |
QY 243 TGTCAAAAAAGCGTGTGTGACATGATGAAGCCATTTCTCAATCAATCACTGATCACTGA 302
| | | | |
Db 10972 TGTCAAAAAAGCGTGTGTGACATGATGAAGCCATTTCTCAATCAATCACTGATCACTGA 11031
| | | | |
QY 303 CAAGATAGATCAGATCCTTGAGAGCGCTGGAACGATCGTGAACGCTGAGGACGACCC 362
| | | | |
Db 11032 CAAGATAGATCAGATCCTTGAGAGCGCTGGAACGATCGTGAACGCTGAGGACGACCC 11091
| | | | |
QY 363 CTCTATCTCTGACAGAGGTTGAGAGATCAAGAAACAGATCACTGAATAATGATGTGC 422
| | | | |
Db 11092 CTCTATCTCTGACAGAGGTTGAGAGATCAAGAAACAGATCACTGAATAATGATGTGC 11151
| | | | |
QY 423 AGTAGACATGAGAAAGCTACACCGCTGTGATGAAACTCTTAAACAGAGGGAGAGAAAT 482
| | | | |
Db 11152 AGTAGACATGAGAAAGCTACACCGCTGTGATGAAACTCTTAAACAGAGGGAGAGAAAT 11211
| | | | |
QY 483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCAAGCTGTTCAAGATAAGCT 542
| | | | |
Db 11212 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCAAGCTGTTCAAGATAAGCT 11271
| | | | |
QY 543 TGACCAATGCTTTTCATTTGGGAGAACTACACACACTGTGGAAGAGAGGAGCCAA 602
| | | | |
Db 11272 TGACCAATGCTTTTCATTTGGGAGAACTACACACACTGTGGAAGAGAGGAGCCAA 11331
| | | | |
QY 603 ACTACTGATGATGATGAGACTGACAGAAAGTTCTGTGTGATCAATGTCTATGATAGT 662
| | | | |
Db 11332 ACTACTGATGATGATGAGACTGACAGAAAGTTCTGTGTGATCAATGTCTATGATAGT 11391
| | | | |
QY 663 TACCATTAAGAATCACTCAAGATTTTCATCCGGGACCTGGAAGATCCCTGAATGATCTTC 722
| | | | |
Db 11392 TACCATTAAGAATCACTCAAGATTTTCATCCGGGACCTGGAAGATCCCTGAATGATCTTC 11451
| | | | |
QY 723 AGTAGTAAGAAACACAGCAAGAGAGCAGAGACCATTAAGGAGAAATAGATGACTACA 782
| | | | |
Db 11452 AGTAGTAAGAAACACAGCAAGAGAGCAGAGACCATTAAGGAGAAATAGATGACTACA 11511
| | | | |
QY 783 GAGAGAGCTGATATGATTAATTAACCTTAGGTTCTGAACATCATTTGGCGCATGTGGGAGCC 842
| | | | |
Db 11512 GAGAGAGCTGATATGATTAATTAACCTTAGGTTCTGAACATCATTTGGCGCATGTGGGAGCC 11571
| | | | |
QY 843 TGATTAAGCCATTTGCAAGAGATATACATAGTTAAATTCAGGATGGGATTCCTTA 902
| | | | |
Db 11572 TGATTAAGCCATTTGCAAGAGATATACATAGTTAAATTCAGGATGGGATTCCTTA 11630
| | | | |
QY 903 ATTAAGCTTGAAGAGCCGATTTGACCAATTGAG-----AGCAATGACGCTGCCGTCA 956
| | | | |

Db 11631 ATAAAGCTTGGAAAGCCGATGTGCAAACTTGAAGAGCAATGACGGCTGCCCTTCACT 11690
QY 957 GTACAGATGAGACTGACGGCGTATTTTCTGGCTGATATTTGACAGGTGGTAACTTCCCTT 1016
Db 11691 ACCAGGATGAGCTGACGGCGTATTTTCTGGCTGATATTTGACAGGTGGTAACTTCCCTT 11750
QY 1017 CAATGCTCCAAATTTGGAACAGCATCTGCAACAGTGTGCAAGATTTGATGTGTAGAG 1076
Db 11751 CAATGCTCCAAATTTGGAACAGCATCTGCAACAGTGTGCAAGATTTGATGTGTAGAG 11809
QY 1077 CAATTTAAGTGTAGGCGCTATCAACAGCAGATAG-AGTGTATGCACTGACTCATCAACCA 1135
Db 11810 CAATTTAAGTGTAGGCGCTATCAACAGCAGATAGAAATGGAAGAAGCTGAATCATCAACCA 11869
QY 1136 GACGCTTTTGTCTAAAGAAAGTACAGNAGAGTGAACAACACTGTTCTTGAAGCCCAA 1195
Db 11870 GA-CCCTTTGTCTAAAGAAAGTACAGNAGAGTGAACAACACTGTTCTTGAAGCCCAA 11928
QY 1196 AATGGANCTGATATTTGATAT-GGCTAGGCTGGAGNAGAAATCATCAACAGCAGCATTA 1254
Db 11929 AATGGANCTGATATTTGATATGAGATAGCTTGGAGAGAAATCATCAACAGCAGCATTA 11988
QY 1255 ACTGAGGCTGCTCTATAGCCCTGAGTCACTGCTGATGAGCTCTG-- 1312
Db 11989 ACTGAGGCTGCTCTATAGCCCTGAGTCACTGCTGATGAGCTCTG-- 12048
QY 1313 CATGCTGACACACACCGAGGCTGCTTAAGTGAAGCAAACTGTTGGAGAGACCTTAA 1372
Db 12049 ATGCTGACACACACCGAGGCTGCTTAAGTGAAGCAAACTGTTGGAGAGACCTTAA 12108
QY 1373 AGCCATTGAAATTTGAACCTTGCACAGCATCATGTGCTCCAAATATGATTAAGCCCATCA 1432
Db 12109 AGCCATTGAAATTTGAACCTTGCACAGCATCATGTGCTCCAAATATGATTAAGCCCATCA 12168
QY 1433 GTCCACAGTGAAGCCGTTAAATGAAGCAGAAATGATCTAAATGATCAAGTGAAGAGA 1492
Db 12169 GTCCACAGTGAAGCCGTTAAATGAAGCAGAAATGATCTAAATGATCAAGTGAAGAGA 12228
QY 1493 AGAAGCAAGCACTTCAAGAACAGTGAAGGTTTAAATCAACGCTGGCAAAATGTTTT 1552
Db 12229 AGAAGCAAGCACTTCAAGAACAGTGAAGGTTTAAATCAACGCTGGCAAAATGTTTT 12288
QY 1553 GGAAGAAACAGAACAAAGAGAGCAGCAGTGAAGTGTGCTTGGCCAGGCGCAAGGCTT 1612
Db 12289 GGAAGAAACAGAACAAAGAGAGCAGCAGTGAAGTGTGCTTGGCCAGGCGCAAGGCTT 12348
QY 1613 CCATGCGCAAAATTTGAGATTTGACAGCAGTGTGCTGACACAGAGCTCATCTGTTGGC 1672
Db 12349 CCATGCGCAAAATTTGAGATTTGACAGCAGTGTGCTGACACAGAGCTCATCTGTTGGC 12408
QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGTCCATATGA 1732
Db 12409 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGCAGCTTAATGTCCATATGA 12468
QY 1733 AGTGTGCTGCTCTTGAAGCTTAAAGAAACAATTAAGATCTGATGCGAAAGGCCA 1792
Db 12469 AGTGTGCTGCTCTTGAAGCTTAAAGAAACAATTAAGATCTGATGCGAAAGGCCA 12528
QY 1793 GCAGATGCTTGCAGATATCCCAAAATCTGCAGAGCAAAATTTGACCAAGACATTAATTA 1852
Db 12529 GCAGATGCTTGCAGATATCCCAAAATCTGCAGAGCAAAATTTGACCAAGACATTAATTA 12588
QY 1853 CTTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1887
Db 12589 CTTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 12623

RESULT 4
US-10-108-260A-1453
; Sequence 1453, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1453
; LENGTH: 3307
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1453
Query Match 80.6%; Score 1637.6; DB 16; Length 3307;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1778; Conservative 0; Mismatches 58; Indels 18; Gaps 9;
QY 8 GCTGACAGAAACACATCAATCATCTCTGACG-TCCGCGCCAGCCCTTGAATATGATAC 66
Db 1457 GCTGACAGAAACACATCAATCATCTCTGACGTTCCGCGCCAGCCCTTGAATATGAAAC 1516
QY 67 TCTAAGCAGCAGCAGAGAAACATTCGCGCAATCGCTGATGATTGATAGCTGAACACAGCC 126
Db 1517 TCTAAGCAGCAGCAGAGAAACATTCGCGCAATCGCTGATGATTGATAGCTGAACACAGCC 1576
QY 127 TCAATATGATTAAGTGAACAAACTGG---CCAAGATCTGAATGAGCCCTGGGGAAG 182
Db 1577 TCAATATGATTAAGTGAACAAACTGGGCGCACAGTACTGGAATTTGAGCCCTGGGGAAG 1636
QY 183 CTTTCTTATCCAAAGAGATATGTGGACCGGACACCCCTTACAGTCAATTAAGAAGA 242
Db 1637 CTTTCTTATCCAAAGAGATATGTGGACCGGACACCCCTTACAGTCAATTAAGAAGA 1696
QY 243 TGTCAAAAGCGTCTGTGCACTGATGAAGCAATTTCTCATCAACTCAGTTTCATGA 302
Db 1697 TGTCAAAAGCGTCTGTGCACTGATGAAGCAATTTCTCATCAACTCAGTTTCATGA 1756
QY 303 CAAGATGATCAGATCCTTGAGAGCCGGAACGATGCTGGAAGTCTGAGGCGCACCC 362
Db 1757 CAAGATGATCAGATCCTTGAGAGCCGGAACGATGCTGGAAGTCTGAGGCGCACCC 1816
QY 363 CTTATCTCTGACAGGTTGGAAGATCAAGAAACAGATCAGTGAATTAAGAATGTGTC 422
Db 1817 CTTATCTCTGACAGGTTGGAAGATCAAGAAACAGATCAGTGAATTAAGAATGTGTC 1876
QY 423 AGTAGACATGAAAGCTACAGCCGTTGATGAACCTTTAAACAGAGGAGAGAAAT 482
Db 1877 AGTAGACATGAAAGCTACAGCCGTTGATGAACCTTTAAACAGAGGAGAGAAAT 1936
QY 483 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCGCAAGCTGTTCAGATTAAGCT 542
Db 1937 GATTGCTAGATCTGGGGGAGCTGATTAAGACATATCTGCGCAAGCTGTTCAGATTAAGCT 1996
QY 543 TGACCAATAGGTTTCTATTTGGGAGACATACACACTGCTGGAAGAGAGGAGCA 602
Db 1997 TGACCAATAGGTTTCTATTTGGGAGACATACACACTGCTGGAAGAGAGGAGCA 2056
QY 603 ACTACGATGTGTAGAGCTAGCAGAAAGTTCTGTGTGATCACAATGCTATGATAGT 662
Db 2057 ACTACGATGTGTAGAGCTAGCAGAAAGTTCTGTGTGATCACAATGCTATGATAGT 2116
QY 663 TACCATTAAGATATCTCAAGTTTCAATCCGGGACCTGGAAGATCTGGAATTTGATCTTC 722
Db 2117 TACCATTAAGATATCTCAAGTTTCAATCCGGGACCTGGAAGATCTGGAATTTGATCTTC 2176
QY 723 AGTAGTAAACCAACAGAGAGAGAGACCATTAAGGAAGAAATTAAGTGAATCA 782
Db 2177 AGTAGTAAACCAACAGAGAGAGAGACCATTAAGGAAGAAATTAAGTGAATCA 2236
QY 783 GGAGAGCTGATATTAATTAACTTAGGTTCTGAATCTATTCGGCATGTGGGAGCC 842
Db 2237 GGAGAGCTGATATTAATTAACTTAGGTTCTGAATCTATTCGGCATGTGGGAGCC 2296
QY 843 TGATAAACCATTTGTCAAGAGATATACATGATTAATTTCAAGCATGGGATTTCTTAA 902

Db 2297 TGATTAACCATTCATGTAAGAGATTAAGTAAATTCAGC-ATGGGATTCCTTAA 2355
QY ATAAAGCTTGGAAAGACCGGATTGACCAACTGTGAG-----AGCAATGACAGCTGCCGTCA 956
Db 2356 ATAAAGCTTGGAAAGACCGGATTGACCAACTGTGAGAGGCAATGACAGGCTGCCGTCACT 2415
QY 957 GTACAGATGACCTGACAGGCGGATTTTCTGGGTAGATATTGACAGGTGTAGTTCGCTT 1016
Db 2416 ACCAGAGTGAAGCTGACAGGCGGATTTTGACTGGGTAGATATTGACAGGGTGTAGTTCGCTT 2475
QY 1017 CAATGCTTCAATTTGGAAAGCATCTGAAACTGTCAAGACAGCATGATGTGCTTAGAG 1076
Db 2476 CAATGCTTCAATTTGGAAAGCATCTGAAACTGTCAAGACAGCATGATGTGAAAGGCTTAAG 2534
QY 1077 CAATTTAAGCTGAGGCGCTATCAACAGCAGATAG-AGTGTAGAGCTGATCATCAAGCA 1135
Db 2535 CAATTTAAGCTGAGGCGCTATCAACAGCAGATAGAAATGAAAGACTGAATCATCAAGCA 2594
QY 1136 GACGCTTTTCTAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTCTAGACCCAA 1195
Db 2595 GA-GCTTTTCTAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTCAAGCCCAT 2653
QY 1196 AATGAGATCTAATTTGATAT-GGCTAGCCTGAGNAGATCATCAACAGACATTA 1254
Db 2654 AATGAGATCTAATTTGATATGAGATAGCTGAGAGAGATCATCAACAGACATTA 2713
QY 1255 ACTGAGGAGGCTCTATTTAGCCNTGGGTGAGTTCACACATGCCCTGTATAGGCTCCG-- 1312
Db 2714 ACTGAGGAGGCTCTATTTAGCCNTGGGTGAGTTCACACATGCCCTGTATAGGCTCCG 2773
QY 1313 CATGCTGACACACACCGAGGCGTCTTAAGTGAACAGAAACCTGTTGAGAGAGACCTTA 1372
Db 2774 ATGGCTGACACACACCGAGGCGTCTTAAGTGAACAGAAACCTGTTGAGAGAGACCTTA 2833
QY 1373 AGCATTGGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCA 1432
Db 2834 AGCATTGGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATGATATTTAGCCATCA 2893
QY 1433 GTCACAGTGGAGCGCTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGA 1492
Db 2894 GTCACAGTGGAGCGCTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGA 2953
QY 1493 AGAAGCAAGCAACCTTCAGAAACAAGCTAAGAGTTTAAATCAAAGCTGCGCAAAATGTTT 1552
Db 2954 AGAAGCAAGCAACCTTCAGAAACAAGCTAAGAGTTTAAATCAAAGCTGCGCAAAATGTTT 3013
QY 1553 GGAAGAAACAGAACAAAG 1612
Db 3014 GGAAGAAACAGAACAAAG 3073
QY 1613 CCATGCGGAATTTGAGAGATTTGACAGAGTGGCTGACTGACACGAGAGGCTATGTTGGC 1672
Db 3074 CCATGCGGAATTTGAGAGATTTGACAGAGTGGCTGACTGACACGAGAGGCTATGTTGGC 3133
QY 1673 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGACAGCTTAATGCTCATATGA 1732
Db 3134 ATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAGAGACAGCTTAATGCTCATATGA 3193
QY 1733 AGTCTGTGCTGCTTTGAGAGACTTAAGAAACAATATAGAGTCTGATGAGAAAGGCCA 1792
Db 3194 AGTCTGTGCTGCTTTGAGAGACTTAAGAAACAATATAGAGTCTGATGAGAAAGGCCA 3253
QY 1793 GCAGATGCTTGAAGATGCCCAAAATCTGACAGAGCAAAATATTGACCAAGACAT 1846
Db 3254 GCAGATGCTTGAAGATGCCCAAAATCTGACAGAGCAAAATATTGACCAAGACAT 3307

APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steimann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12468
LENGTH: 3987
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-12468

Query Match 39.7%; Score 807; DB 14; Length 3987;
Best Local Similarity 95.1%; Pred. No. 5.2e-200;
Matches 884; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

QY 962 GATGACTGACAGCGGATTTTCTGGGTAGATATTGACAGGTGATAGTTCGTTCAATG 1021
Db 46 GATGACTGACAGCGGATTTTCTGGGTAGATATTGACAGGTGATAGTTCGTTCAATG 105
QY 1022 TCTCCAAATTTGGAACAGCATCTGAAAACCTGCAAGCAGCATTTGATGTCTAGACAA 1081
Db 106 TCTCCAAATTTGGAACAGCATCTGAAAACCTGCAAGCAGCATTTGATAGACTTAAGCA 164
QY 1082 TAACTGAGGCGCTATCAACAGAGATAG-AGTGTAGACAGTCACTACAGCAGACGC 1140
Db 165 TAACTGAGGCGCTATCAACAGAGATAGAAATGAAACATGATCATCAAGCAGAG-GC 223
QY 1141 TTTTGTAAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTTGAAGCCCAAAATG 1200
Db 224 TTTTGTAAAGAAAGTAAACAGNAGAGTGAACAAACACACTGTTTGAAGCCCAAAATG 283
QY 1201 AACTGATATTGATAT-GGCTAGCCTGAGAGAGAGAGATCATCAACAGACATTAAC 1259
Db 284 AACTGATATTGATATTGAGATAGCTGAGAGAGAGATCATCAACAGACATTAACCTGG 343
QY 1260 AGGCTCTCTATTAGCGGCTGAGTTCACATGSCCCTGTATGAGTCTCTG--CATGC 1317
Db 344 AGGCTCTCTATTAGCGGCTGAGTTCACATGSCCCTGTATGAGTCTCTGCGATGGC 403
QY 1318 TGAACACACCGAGGCGTCTTAAGTGAACAGAAACCTGTTGAGAGAGACCTTAAGCCA 1377
Db 404 TGAACACACCGAGGCGTCTTAAGTGAACAGAAACCTGTTGAGAGAGACCTTAAGCCA 463
QY 1378 TTGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATATATGATTTAGCCATCA 1437
Db 464 TTGAATTTGAATCTTGCCAGACATCATGTGCTCCAAATATATGATTTAGCCATCA 523
QY 1438 CAGTGAAGCGCTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAG 1497
Db 524 CAGTGAAGCGCTTAATTAAGAGAGAAATGATCTAATTTGAATCAAGTCAAGAGAGAG 583
QY 1498 CAAGCAACCTTCAGAACCAAGCTAAGAGTTTAAATCAACGCTGCGCAAAATGTTTGGAAA 1557
Db 584 CAAGCAACCTTCAGAACCAAGCTAAGAGTTTAAATCAACGCTGCGCAAAATGTTTGGAAA 643
QY 1558 AAAACAGAACAAAG 1617
Db 644 AAAACAGAACAAAG 703
QY 1618 GCGAAATTTGAGATTTGACAGAGTGGCTGACAGACGAGAGGCTCATCTGTTGGCATTTA 1677
Db 704 GCGAAATTTGAGATTTGACAGAGTGGCTGACAGACGAGAGGCTCATCTGTTGGCATTTA 763
QY 1678 AACCGTGGAGGTTTAAACCGGAACAGCCAAAGAGAGCTTAATGCTCATATGAAAGTCT 1737

|||||
Db AACGCTGGAGGTTTACCGGAAACAGCAAGAGACAGCTTATATGCAATGGAAGCT 823
Oy 1738 GTGCTGCTTTGAAAGCTAAAGAAACATATTAAGTCTGATGCAAGAAAGCCGACGA 1797
Db 824 GTGCTGCTTTGAAAGCTAAAGAAACATATTAAGTCTGATGCAAGAAAGCCGACGA 883
Oy 1798 TGCTTGAAGATGCCCAAAATCTGAGAGACAAATATTGACCAAGACATTAATACTTGA 1857
Db 884 TGCTTGAAGATGCCCAAAATCTGAGAGACAAATATTGACCAAGACATTAATACTTGA 943
Oy 1858 AAAAAAAAAAAAAAAAAAAAAAAAAAANAC 1887
Db 944 AAGAAAAATGGGAATCGGTGGAACCAAC 973

RESULT 6
US-10-198-846-12599/c
; Sequence 12599, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhilber, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12599
; LENGTH: 18529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-12599

Query Match 32.1%; Score 652.2; DB 14; Length 18529;
Best Local Similarity 62.8%; Pred. No. 3.7e-159;
Matches 1132; Conservative 0; Mismatches 647; Indels 24; Gaps 7;

Oy 41 CCCCCCCCCCTTTGAAATATGATCTTAAGGCGACGAGGAACATCGCACTG 100
Db 6165 CCCTCTCCAGCATGATCATGAGCAGCTCAGGCGACAAAGAGAAATGAGCAATTA 6106
Oy 101 CGTAGTGTAGTGAAGTGAACAGAGCTCATATGATTAAGTGAACAAACTGGCC---A 156
Db 6105 AGGGAATTAATGCTGAACACAACTTAATTAAGCACTTAATAAGTGGCCACAA 6046
Oy 157 CAGTACTGAATGAGCCCTGGGAAAGGCTTTCTATCCAAGAAATATGTCGACGCCAG 216
Db 6045 CTAAGAAATTAACCTGAGAGAGGGAATGCTGAAGAAATACCAAGAAACAGAA 5986
Oy 217 ACCCTTAACGTCACTTAAGAAAGATGTCAAAAAGCGTGTGGCACTGGATGAAGCC 276
Db 5985 AACATGTATGCCCAATTAAGAGAGAGGTGGCGCCAGCAGCCCTGGCTCTGGATGAAGCC 5926
Oy 277 ATTTCTCAATCAATC-----AGTTCCATGACAGATGATGATCAATCTTGAAGC 327
Db 5925 GTGTCCAGTCCACACAGATTACAGAGTTTCATGATTAATTAAGTGAAGTGAAGCA 5866
Oy 328 CTGGAACCATCTGTGAAGCTGTGAGGAGCAACCTCTATCTCTGAGAGGTTGAAG 387
Db 5865 CTGGAAGATCTTCTCTCGCTGCTGTATGCAACACATGATCTCTGCTGAAGTGAAG 5806
Oy 388 ATCAAGAAACAGATGAGTGAATTAAGATGTGATGATGACATGGAAGCTACAGCCG 447
Db 5805 ATCAAGAGTGCATGATGAGTGAATTAAGATGTGATGATGAGTGAAGAACTGAGCA 5746

Oy 448 TTGATGAACTTTAAACAGAGGAGAGAGAAATGATGCTAGATCTGGGGAGCTGAT 507
Db 5745 TCCTTTGAGGCTTTGAAGCCGCTGGAGAGAGCTTAATGGAACATCTCAGGAGCAGAC 5686
Oy 508 AAAGCATATCTGCCAAAGCTTTCAAGATTAAGCTTGAACAAATGTTTCAATTTGGAG 567
Db 5685 AAGATCTGCTGCAAAAAGAAATCAGAGTAAATTTGATCAAAATGTTTCTTCTGGAG 5626
Oy 568 AACATACACACTGTGAG 627
Db 5625 GACATCAAAAGCTCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5566
Oy 628 GAAAGTCTGTGTGATGATCAGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db 5565 GAGAGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5506
Oy 688 ATCCGGGACCTGGAAGATCTGGAATGATCTTCAATGATTAACAAACAGAGAGAG 747
Db 5505 GTCCATGACTTGAAGAGCCGAGGATGATCTTCCATCAATCAACAGAGTGAAGCT 5446
Oy 748 GCAGAGACATTAAG 807
Db 5445 GCTGAGACTAATTAAG 5386
Oy 808 CTAGGCTCTGAACCTATGCGGAGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867
Db 5385 CTGAGAGAGATTTGATTTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5326
Oy 868 ATACATGATTAATTAAG 927
Db 5325 ATGATGATGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5267
Oy 928 CCAACTGAG-----AGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
Db 5366 AAAAATTGAG 5207
Oy 982 TTTCTGAGTATTAATTAAG 1041
Db 5206 TGACTGCTGATTAATTAAG 5148
Oy 1042 TCGAAGCTGTCAG 1101
Db 5147 TCAATGATGATTAATTAAG 5088
Oy 1102 AGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
Db 5087 AGCAAAATTAAG 5028
Oy 1162 NAGAGAGTGAACACACAGCTGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
Db 5027 ATGAG 4968
Oy 1221 GCTGAG 1280
Db 4967 ACCTGGGAG 4908
Oy 1281 GTCAATTCACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338
Db 4907 GTCAATTCACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4848
Oy 1339 TAAGTGAAG 1398
Db 4847 TAAGTGAAG 4788
Oy 1399 ATCATGCTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
Db 4787 ACATGCTCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4728
Oy 1459 CAGGAATGATTAATTAAG 1518
Db 4727 CTGCAATGAG 4668
Oy 1519 TAGAGGTTTAAATCAAG 1578

Db 4667 TGGAAACATGAACCAATGCTGGAGTCAAGTGTTCACGAAACAGAGAGAGAGAGC 4608
Qy 1579 AGCTGATGCTGCTTCGCGCCAGCCAAAGAGTTCCATGGCGAATTTGAGATTTGACG 1638
Db 4607 AGCTTCAGTAACTCTGACAGAGCCAGGCTTCCAGATGAATTTGAGATTTCTCT 4548
Qy 1639 AGTGGCTGACTGACAGGAGCGTCAATCTGTGGCATTTAAACCGCTGGAGGTTTACCG 1698
Db 4547 TGGAACTTAATAAGATGAGACGAGCTTCTGCATCTAAGCCCAAGAGAGCTTCTG 4488
Qy 1699 AAACAGCCAGAGAGAGCTTAATGTCCATATGGAAGTCTGTGCTTGAAGCTTAAG 1758
Db 4487 AAACCTGTAGGAAACAGCTTGATACATATGGAACCTTAATCCAGCTGAAGCCAGG 4428
Qy 1759 AAGAAACATATAAGAGTCTGATGCAAGAAAGCCAGAGATGCTTGAAGATGCCAAAT 1818
Db 4427 AAGAGACTTAATATCAACTTCTGACAGAGGCGAGACTCATGCTTTAGCCGTGACGACT 4368
Qy 1819 CTG 1821
Db 4367 CTG 4365

RESULT 7
US-09-823-245A-445
Sequence 445, Application US/09823245A
Publication No. US20020039760A1
GENERAL INFORMATION:
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakar
APPLICANT: Graham, James R.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6401
CURRENT APPLICATION NUMBER: US/09/823,245A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/194,941
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 445
LENGTH: 2181
TYPE: DNA
ORGANISM: Homo sapiens
US-09-823-245A-445

Query Match 24.7%; Score 501; DB 9; Length 2181;
Best Local Similarity 62.0%; Pred. No. 4,4e-120; Indels 19; Gaps 6;
Matches 895; Conservative 0; Mismatches 529;

Qy 296 TCCATGACAGATAGATGATCAGATCCTTGAGAGCTTGAGACCATCTGGAAGCTCTGAGC 355
Db 1 TCCATGATAAATTTGAGCCATATGTTGAGAGACACTGAGAGATCTTCTCCTCCGCTGGTA 60
Qy 356 AGCCACCCCTTATCTCTGACAGAGTTGAGAGATCAAGAGAACAGATCAGTGAATAAGA 415
Db 61 TGCACCACTGATCCCTGCTGAGAGTAGAACAGATCAGAGAGTGCATCAGTGAACATTAAGA 120
Qy 416 ATGTGTCAGTACATGAGAAAGCTACAGCCGTTGATGAACTCTTTAAACAAGGGAG 475
Db 121 GTGCCACCGTGAAGCTAGAAAACCTGACGCTATCTTTGAGCCCTTGAAGCGCCGTGAG 180
Qy 476 AGGAAATGATTTGATAGTCTGAGGAGCTGATGAAGATATCTGCCAAAGCTTTCCG 535
Db 181 AGGAGCTTATTTGAGAGATCTCAGGAGACAGAAAGATCTGGCTGCAAAAAGAAATCCAG 240
Qy 536 ATAGCTTGACCAATGCTTTTCAATTTGGAGAAACATTAACACACTGGTGAAAGAGAGG 595

Db 241 ATAAATGATCAAAATGATATCTCTGGAGAGACATCAAGCTCGGCGTGAAGACGAG 300
Qy 596 AAGCCAACTACTGATGATGATGAGAGCTAGCAAGAAAGTTCTGGTGTGATCATATCAT 655
Db 301 AAATCAAAATTTCTTGATGCTCTTGAATTTGACAGAAAGTTCTGGTGTGATGATGAG 360
Qy 656 TGAATGTTACATTAAGATTAATCAAGATTTGATCCGAGACCTGGAAGATCTCGAATTG 715
Db 361 TCTGACCAACATCAAAAGACCCAGGATATGTCATGATGATCTGGAAGCCAGGACTTG 420
Qy 716 ATCTTCACTAGTAAATCAACAGCAAGAAAGCAGCAGAGCATTAAGGAAAGAAATGATG 775
Db 421 ATCTTCCATCATCAAAACAAAGGTTGAAGCTGCTGAGACTTAAGAAAGACAGATG 480
Qy 776 GACTACAGAGAGAGCTGATATGATTAATTAACCTAGCTTCTGAACCTATGGGCGCATG 835
Db 481 GTCTGATBAAGAGCTGAGAGTTATTCGATCTTGGAGCAGATTTGATTTTGGCTGTG 540
Qy 836 GAGAGCTGATTAACCCATTTGTCAAGAGATATCATGATTAATTAATTCAGCGATGGAT 895
Db 541 GAGAACTGAGAAAGCTGAAATGAATTAATGCTTGGAGAACTTAACGGAACATGG--- 597
Qy 896 TCTTAATTAAGCTTGGAAAGACCGGATTAACAACTTGAAGCAATGCACTGCGCTC 955
Db 598 ---AAGATAGGCTAGCAAAAATTG-----AAGATGCTATGCAAGCTGCTGTGCA 646
Qy 956 AGTACAGATGAGACTGACAGCGGATTTTCTGGGTGATATATTCAGGTGTAAGTTCGCT 1015
Db 647 TATCAGAGAGCTTTTACGCTATGTTGACTGACAGATTAACCTGATTAACCTTTC 706
Qy 1016 TCAATGTCTCAATTTGGAACAGCATCTCGAACTGTCAAGCAGCAGATTAATGCTAG 1075
Db 707 ACCAAGCCCCCTGTTGGCTGTG-ACCTGATCTGTTAAGATCACTTAATGAATGA 765
Qy 1076 GCAATTTAAGTCTGAGGCTTATCAACAGCAGATAGTGTGATCAGTCACTATCAGCA 1135
Db 766 GGAGTTCAAGTGAAGGTATACAGAGCAAAATGAGAGAACCTTAATCACCAGG 825
Qy 1136 GACGCTTTTGTAAAGAAAGTAAAGAGAGTGAACAAACACAGTCTTGAACCCAA 1195
Db 826 TGAATGATGTTAAAGAAAGCTATGATGAGACGCAAGACATTAATACAGAACTACT 885
Qy 1196 AATGAGCTGATATGATATGCT-AGCTGAGNAGAGATCATCAAGACAGCATTA 1254
Db 886 GACAGAACTCAAAACCTCTGGGAGAACCTGGGTGAAGAAATGGCCACGACAGCA 945
Qy 1255 ACTGAGGGTGTCTATTAGCCNTGGTCAATTCACATGCCCTGTATGAGTCTGTCA 1314
Db 946 ACTAGAAAGGGCTCTGTTGGCCCTTGTGATGATGCTGATGCTTGAAGAACTAATGAG 1005
Qy 1315 T--GCTGACACACACGAGGGCTTGTAGTGAAGAAACGTTGAGAGAGACCTTA 1372
Db 1006 TTGGCTGACTCATACCGAAGAGTTGTTAGATGCTCAGAAACAAATTAAGTGAACCCAA 1065
Qy 1373 AGCCATTTGAATTTGAATCTTCCAGCATGATGCTCCAAATGATGATTAAGCCCATCA 1432
Db 1066 AGTCATTTGAATTTGAGCTGCAAGACCATGTCCTTAAGAAATGATGTTTGGCTATCA 1125
Qy 1433 GTCCACAGTGAAGCCGTTAATTAAGCAGAAATGATCTTAATGAATCAAGTGCAGAGA 1492
Db 1126 AGCCACAGTGAAGAACGTCAACAAAGCTGGCAATGAGCTTCTTGAATCCAGTCTGAGA 1185
Qy 1493 AGAAGCAACCACTTCAAGCAAGTGAAGTTTAATCAACGCTGCAAAATGTTT 1552
Db 1186 TGAATGCAAGCTTAAAGAGCCGTTTGAAGCAATGAACCAATGCTGAGATCAATGTT 1245
Qy 1553 GGAAGAAACAGAAACAGAGACAGAGCTGATGCTTGGCCAGCAAGAGGTT 1612
Db 1246 ACAGAAACAGAGAGAGAGAGAGAGCTTCAATCTGTCAGACAGCCAGGCTT 1305
Qy 1613 CCATGGCGAAATTTGAGATTTGACAGAGTGGCTGACTGACAGAGCGTATCTGTTGGC 1672

Db 1306 CCACAGTGAATGTAAGA-TTCCTCTTGGAATTACTAGAAATGAGAGCAGCTTTCTGC 1364
Qy 1673 ATCTAAACCGCTGGAGGTTTACCGGAACCGCAAGAGAGCTTAATGTCATATGGA 1732
Db 1365 ATCTAAACCGCAAGAGAGACTTCTGAAATGCTAGGAGAACGCTTGATACATATAGT 1424
Qy 1733 AGT 1735
Db 1425 AAT 1427

RESULT 8
US-10-158-034-52
; Sequence 52, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; PRIORITY FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 2113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-158-034-52

Query Match 24.5%; Score 497.8; DB 15; Length 2113;
Best Local Similarity 63.7%; Pred. No. 3e-119;
Matches 836; Conservative 0; Mismatches 466; Indels 11; Gaps 5;

Qy 518 CTGCCAAAGCTGTTCAGGATAGCTTGACCAATGCTTTTCATTTGGAGAACTACACA 577
Db 62 CGGCAAAAAGAAATCCAGGATAAATTCATCAATGCTATTTCTTGGAGAGACATCAAG 121
Qy 578 CACTGCTGGAAGAAGAGGAGCCAACTACTGATGATGATGAGCTAGCAGAAAATTCT 637
Db 122 CTCGGGCTGAAGAACGAAATCAATTTCTGATGCTCTTGAATAGCAGAGAACTTCT 181
Qy 638 GGTGTGATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 697
Db 182 GGTATGACATGCGACCTCTCTGACACCATCAAGAGACCCAGATATTTGCTGACT 241
Qy 698 TGAAGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 757
Db 242 TGAAGAGCCAGGACTTGAATCTTCTGATGATGATGATGATGATGATGATGATGAT 301
Qy 758 TAAGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 817
Db 302 TTAAGGAAGAGACAGATGCTGATGATGATGATGATGATGATGATGATGATGATG 361
Qy 818 AACCATTTGGCGGAGTGGGAGGCTGATAAACCATTGTAAGAGATATACATGAGT 877
Db 362 ATTGATTTTTCCTGAGAGAACTGGAAGCCGTAAGTAGAGAGATGATGATGATG 421
Qy 878 TAAATTCAGCATGAGATTTCTTAATTAAGCTTGAAGAGCCGATGACCACTTGAG 937
Db 422 TGAATATATGC-TTGGAGAACTTAAACAAACATGGAAGAGAGGCTTGAAGAACTTG 480
Qy 938 -----AGCAATGACGCTCCGCTCAGTACAGATGAGCTGACGGCGGATATTTTCTGG 991
Db 481 GATGCTATGCAAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 992 GATATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1051
Db 541 GATTAACATGATGATTAATTAATCTGACACATGCCCCCTGTGGACCTG-ACCTCAATAC 599
Qy 1052 CAACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1111
Db 600 TAAAGATAGTTAATGAATGAAGAGCTTCAAAGTATGAATTTTCAACAGCAAAATTGA 659

Qy 1112 GTGTAGAGCTGACTCATACGAGAGCCTTTTGTCTAAAGAAAGTAAACAGNAGAGTGA 1171
Db 660 GATGAGAGAGCTTAATATACAGAGGCTGATGATGATGATGATGATGATGATGATGATG 719
Qy 1172 CAACACACTGTTCTAGAACCCAAATAAGATGATGATGATGATGATGATGATGATGATG 1230
Db 720 CAGAGACTTATATACAGAACCACTGACAGAACTCAACCTCTGGAGAAACCTGGGCTGA 779
Qy 1231 GAGATCATCAACAGACAGATTAATCTGAGAGGCTCTCTATTTAGCNCATGGCTAGTTCCA 1290
Db 780 GAAATATGCCACCGAGACAGCAAACTAGAAAGGGGCTCTGTGGCCCTTGGTCACTTCCA 839
Qy 1291 ACATGCCCTGTATGAGCTCTGCTGAT--GCTGACACACACCGAGGGCTGTGATGATGAGCA 1348
Db 840 GCATGCCCTTGAAGAGAACTTAATGATGATGATGATGATGATGATGATGATGATGATG 899
Qy 1349 GAAACCTGTTGAGAGAGACCTTAAGCCATTTGAATTTGAATTTGCAAGCATCATGCT 1408
Db 900 GAGACCAATAGTGAAGAGACCCAAAGCTATTGAAGTTGAGCTGCCAAGCAGCATGCTCT 959
Qy 1409 CCAAAATGATGATTTAGCCCATCATGTCACAGTGAAGCCGTTATTAAGCAGAGAAATGA 1468
Db 960 AAAAATGATGATTTTGGCTCATCAAGCCACAGTGAAGAAAGTCAAAAGCTGGCAATGA 1019
Qy 1469 TCTAATTTGATTCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1528
Db 1020 GCTTCTTGAATCCAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1079
Qy 1529 AAATCAAGCTGTGCAAAATGTTTGGAAAAACAGAACCAAGAGAGAGAGAGAGAGAG 1588
Db 1080 GAACCAATGCTGGAGAGTCAATGTTTACAGAAAACAGAGAGAGAGAGAGAGAGAGAG 1139
Qy 1589 TGCTTGGCCGAGGCCAAAGGGTTTCATGCGAAATTTAGAGATTTGACAGAGTGGAC 1648
Db 1140 AACTCTGACAGAGGCCAGAGGCTTCCACAGTGAATTTGAAGATTTCTCTTGGAACTTAC 1199
Qy 1649 TGACACGAGAGCTGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATG 1708
Db 1200 TAGAATGAGAGCCAGCTTCTGATGATGATGATGATGATGATGATGATGATGATG 1259
Qy 1709 GAGACAGCTTAATGTCATATGAGATGATGATGATGATGATGATGATGATGATGATG 1768
Db 1260 GGAACAGCTTGTATACATATGAGATGATGATGATGATGATGATGATGATGATGATG 1319
Qy 1769 TAAGATCTGATGAGAGAGAGCCAGAGATGCTTTCAGATGATGATGATGATGATG 1821
Db 1320 TAATCACTACTTGAACAAGGAGAGAGATGATGATGATGATGATGATGATGATGATG 1372

RESULT 9
US-10-158-034-32
; Sequence 32, Application US/10158034
; Publication No. US20030219758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT210C1
; CURRENT APPLICATION NUMBER: US/10/158,034
; PRIORITY FILING DATE: 2002-05-31
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 5453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-158-034-32

Query Match 24.5%; Score 497.4; DB 15; Length 5453;
Best Local Similarity 63.7%; Pred. No. 6.2e-119;
Matches 835; Conservative 0; Mismatches 465; Indels 11; Gaps 5;

QY 520 GCCAAGCTGTTGAGTAAGCTTGAACCAATGGTTTTCATTTGGAGAAACATACACA 579
 Db 1 GCAAAAAGAAATCCAGAGTAATGATCAATGGTATTTCTTGGGAGGACATCAAGCT 60
 QY 580 CTGTGTGAAGAGGAGGAGCAAACTACTGATGTGATGAGCTAGAGAAAGTTCTGG 639
 Db 61 CGGGCTGAAGAGAGAAATCAAAATTTCTGATGTCTTGAATTTAGAGAGAAAGTTCTGG 120
 QY 640 TGTGATCACATGTCATTTGATGATTTACCAATTAAGATCTCAAGATTTTCATCCGGGACCTG 699
 Db 121 TATGACATGGCAGCTCTCTGACACCATCAAGAGACCCAGATATTTGTCATGACTTG 180
 QY 700 GAAGATCTGGAAATGATCTCTGATGATTAACCAACAGAGAGAGAGAGAGAGACATA 759
 Db 181 GAAAGCCCAAGGACATGATCTTCCATCATCAACACAGGTTGAAGCTGCTGAGACTAT 240
 QY 760 AGGAGAGAAATTAATGACTACAGAGAGAGCTGATTAATTTAACTTGAAGTTCTGAA 819
 Db 241 AAGGAGAGAGCAATGCTGATGAGAGAGCTGAGATTTATTCGATCTTGGAGCAGAT 300
 QY 820 CTGATTCGGGATGTTGGGAGGAGCTGATTAACCCATTTGTCAAGAAAGATATACATGATTA 879
 Db 301 TTGATTTTTCCTGTGAGAGAACTGAGAGCTGAGAGTGAGAGACATTTGATGAGATG 360
 QY 880 AATTCAGCAGTGGGATTTCTTAAATAAAGCTTGAAGAGACCGGATTTGACCACTTGAG-- 937
 Db 361 AATTAATGC--TTGGAGAACTTAAACAAACATGAGAAAGAGGCTGAGAAATCTTGAGGA 419
 QY 938 ----AGCAATGACGCTGCCCTCAGTACAGATGAGCTGACAGGCGGTATTTTCTGGGTAGA 993
 Db 420 TGTATGACAGCTGCTGTGAGATGAGAGCACTCTTCAGGCTATGTTTGAATGAGCTGAGA 479
 QY 994 TATTGAGGAGTGAAGTTCCTTCAATGTCCTCAATTTGGAACAGCATCTCGAAAGCTGCA 1053
 Db 480 TAACTCTGATTAATTAATCTGACACCATGCCCCCTGTTGGACCTG--ACCTTAATACCTGTTA 538
 QY 1054 AGCAGAGATTTGATGATGCTAGAGCAATTAAGTCTGAGGCTTCAACAGCAGATAGAGT 1113
 Db 539 AAGATCAGTTAAATGAATGAAGAGAGTTCAAGTAGAAGTTTACCAACAGCAAAATTGAGA 598
 QY 1114 GGTACGACTGACTCATTCAGCAGACACGCTTTTGTCTAAAGAAAGTAAACAGAGAGAGTACA 1173
 Db 599 TGGAGAGGCTTAAATCAACAGGAGTGAATGATGTTAAAGAAAGGCTACTGATGAGACGAGCA 658
 QY 1174 AACACACTGTTCTAGACCCCAAAATGAGACTGATTTGATAT--GGCTAGCGCTGAGAGAGA 1232
 Db 659 GAGACATTAATACAGAGAACCACTGACAGAACTCAAAACCTCTGGAGAACTGGTGAGAGA 718
 QY 1233 GAATCATCAACAGACAGATTAACCTGAGAGGCTCTTATTAAGCCMTGGGTCAAGTTCCAA 1292
 Db 719 AAATGCCCAACCGACAGCAAAACTAGAGGGGCTGTGTTGGCCCTTGGTCAATTCAGC 778
 QY 1293 ATGCCCTGATAGCTCTCTGAT--GCTGACACACCCAGAGGCTTGTCTAAGTGAAGACA 1350
 Db 779 ATGCCCTTGAAGCACTTAATGATGTTGGCTGACTCATACCGAAGAGTTGTTAGATGCTCA 838
 QY 1351 AACCTTGGAGAGAGACCTTAAGACATTTGAATTTGAATCTTGCACAGCATCATGTGCTC 1410
 Db 839 GACCAATTAAGTGAAGACCCAAAGTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 898
 QY 1411 AAAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
 Db 899 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
 QY 1471 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530
 Db 959 TTTCTTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
 QY 1531 ATCAACGCTGCAAAATGTTTGGAAAAAACAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1590
 Db 1019 ACCAAATGCTGGAGTCAATGTTTACAGAAAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
 QY 1591 CCTTGGCCAGGCAAGAGGTTCCATGCGCAAAATTGAGATTTGACAGCAATGCTGAGCTGAG 1650

Db 1079 CTCTGAGAGAGGCCCCAGGCTTCCACAGTGAATTAAGATTTCTCTTGGAACTTACTA 1138
 QY 1651 ACACGAGGCTCATCTGTTGGATCTTAACCCCTGGAGAGTTTACCGGAAACGCCAAG 1710
 Db 1139 GAATGAGAGGAGGCTTCTGATCTTAAGCCACAGAGAGACTCTTGAATCTGTAGGG 1198
 QY 1711 AGCAGCTTAATGCTCATATGAAAGTGTGCTGCTTGAAGTAAAGAAACATATA 1770
 Db 1199 AACAGCTGATACACTATGAAATCTTATCCAGCTGAAGCAAGAGAGACTTATA 1258
 QY 1771 AGAGTGTATGCAAGAAAGGCCAGAGATGCTTGAAGATGCCCAAAATCTG 1821
 Db 1259 ATCAACTACTTGAACAAGGAGAGACTGATGCTTGAAGCGGTGAGACTCTG 1309

RESULT 10

US-10-198-846-4738
 / Sequence 4738, Application US/10198846
 / Publication No. US2003009974A1
 / GENERAL INFORMATION:
 / APPLICANT: Lillie, James
 / APPLICANT: Xu, Yongyao
 / APPLICANT: Wang, Youzhen
 / APPLICANT: Steilmann, Kathleen
 / TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 / TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 / TITLE OF INVENTION: THERAPY OF BREAST CANCER
 / FILE REFERENCE: MRI-049
 / CURRENT APPLICATION NUMBER: US/10/198,846
 / PRIOR FILING DATE: 2002-07-18
 / PRIOR APPLICATION NUMBER: 60/306,220
 / NUMBER OF SEQ ID NOS: 14064
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 4738
 / LENGTH: 599
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: 599
 / OTHER INFORMATION: n = A,T,C or G
 US-10-198-846-4738

Query Match

Best Local Similarity 21.2%; Score 430.8; DB 14; Length 599;
 Best Local Similarity 94.0%; Pred. No. 5.3e-102;
 Matches 500; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

QY 962 GATGAGCTCAGAGCGGTATTTTCTGGATGATTAATGACAGGTGATGATGATGATGATGATGAT 1021
 Db 61 GATGAGCTCAGAGCGGTATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 1022 TCTCAATTGGAACAGCATCTGGAACCTGCAAGCAGAGATGATGATGATGATGATGATGAT 1081
 Db 121 TCTCAATTGGAACAG-ATCTGGAACCTGCAAGCAGAGATGGAAGACTTAAGCAAT 179
 QY 1082 TAAATGAGGCTTATCAACAGAGATAG-AGTGTGATGATGATGATGATGATGATGATGATGAT 1140
 Db 180 TAAATGAGGCTTATCAACAGAGATAGAAATGGAAGACTGGAATGATGATGATGATGATGATGAT 238
 QY 1141 TTTTGTAAAGAAAGTAAACAGAGAGAGTGAACAAACAACATGTTCTAGACCCCAAAATG 1200
 Db 239 TTTTGTAAAGAAAGTAAACAGAGAGAGTGAACAAACAACATGTTCTAGACCCCAATTAATG 298
 QY 1201 AACTGATATGATAT-GGCTAGCTGAGAGAGATCAATCAACAGACAGATTAATCTG 1259
 Db 299 AACTGATATGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 358
 QY 1260 AGGAGCTTATTAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
 Db 359 AGGAGCTTATTAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418

Oy 1318 TGACACACACCGAGGGCTTGCTAAGTAGCAGAAAACCTGTGGAGGAGACCCCTAAAGCA 1377
Db 419 TGACACACACCGAGGGCTTGCTAAGTAGCAGAAAACCTGTGGAGGAGACCCCTAAAGCA 478
Oy 1378 TTGAAATTTGAATTGCGCAAGCATGTCCTCAAAATGATGATTAAGCCCATGATCA 1437
Db 479 TTGAAATTTGAATTGCGCAAGCATGTCCTCAAAATGATGATTAAGCCCATGATCA 538
Oy 1438 CAGTGAAGCCGTTAATTAAGCAGAAATGATCTTAATTAATCAAGTCAGG 1489
Db 539 CAGTGAAGCCGTTAATTAAGCAGAAATGATCTTAATTAATCAAGTCAGG 590

RESULT 11

US-10-037-270-352
; Sequence 352, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhimei
; APPLICANT: Tillingshaet, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 352
; LENGTH: 4495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (307)..(3006)
US-10-037-270-352

Query Match 13.1%; Score 266.8; DB 15; Length 4495;
Best Local Similarity 93.9%; Pred. No. 1.2e-58;
Matches 277; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1593 TTGGCGCCAGGCAAAAGGTTCCATGGCGAAATTGAGATTGCGAGCAGTGGCTGATGAC 1652
Db 172 TTACTTCAGGCCCAAAAGGTTCCATGGCGAAATTGAGATTGCGAGCAGTGGCTGATGAC 231
Oy 1653 ACGAGCGTCATCTGTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAGAG 1712
Db 232 ACGAGCGTCATCTGTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAGAG 291
Oy 1713 CAGCTTAATGTCATATGGAAGTCTGTGCTGCTTTGAAGCTAAAGAAACATATTAAG 1772
Db 292 CAGCTTAATGTCATATGGAAGTCTGTGCTGCTTTGAAGCTAAAGAAACATATTAAG 351
Oy 1773 AGCTGATGCAAGAAAGCCAGAGATGCTGCGAAGATGCCCAAAATCTGCAGAGCAAT 1832
Db 412 ATTGACCAAGCATTAATTAATCTTAAGAAAGAAATGGGAATGGTGGAACCAAG 466

Db 352 AGTCTGATGCAAGAAAGCCAGAGATGCTTGAAGATGCCCAAAATCTGCAGAGCAAT 411
Oy 1833 ATTGACCAAGCATTAATTAATCTTAAGAAAGAAATGGGAATGGTGGAACCAAG 1887
Db 412 ATTGACCAAGCATTAATTAATCTTAAGAAAGAAATGGGAATGGTGGAACCAAG 466

RESULT 12

US-10-117-722-352
; Sequence 352, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCTP
; CURRENT APPLICATION NUMBER: US/10/117,722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 352
; LENGTH: 4495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (307)..(3006)
US-10-117-722-352

Query Match 13.1%; Score 266.8; DB 15; Length 4495;
Best Local Similarity 93.9%; Pred. No. 1.2e-58;
Matches 277; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 1593 TTGGCGCCAGGCAAAAGGTTCCATGGCGAAATTGAGATTGCGAGCAGTGGCTGATGAC 1652
Db 172 TTACTTCAGGCCCAAAAGGTTCCATGGCGAAATTGAGATTGCGAGCAGTGGCTGATGAC 231
Oy 1653 ACGAGCGTCATCTGTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAGAG 1712
Db 232 ACGAGCGTCATCTGTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAGAG 291
Oy 1713 CAGCTTAATGTCATATGGAAGTCTGTGCTGCTTTGAAGCTAAAGAAACATATTAAG 1772
Db 292 CAGCTTAATGTCATATGGAAGTCTGTGCTGCTTTGAAGCTAAAGAAACATATTAAG 351
Oy 1773 AGTCTGATGCAAGAAAGCCAGAGATGCTGCGAAGATGCCCAAAATCTGCAGAGCAAT 1832
Db 352 AGTCTGATGCAAGAAAGCCAGAGATGCTGCGAAGATGCCCAAAATCTGCAGAGCAAT 411
Oy 1833 ATTGACCAAGCATTAATTAATCTTAAGAAAGAAATGGGAATGGTGGAACCAAG 1887
Db 412 ATTGACCAAGCATTAATTAATCTTAAGAAAGAAATGGGAATGGTGGAACCAAG 466

RESULT 13

US-09-864-408A-111
; Sequence 111, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012

;/ CURRENT APPLICATION NUMBER: US/09/864,408A
;/ CURRENT FILING DATE: 2001-05-24
;/ PRIOR APPLICATION NUMBER: 60/206,690
;/ PRIOR FILING DATE: 2000-05-24
;/ NUMBER OF SEQ ID NOS: 9068
;/ SOFTWARE: FastSeq for Windows Version 4.0
;/ SEQ ID NO 111
;/ LENGTH: 344
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
US-09-864-408A-111

Query Match 12.9%; Score 262.8; DB 11; Length 344;
Best Local Similarity 92.4%; Pred. No. 3.6e-58;
Matches 318; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 1039 ATCTCCAAATCTGTCACGACGACGATGATGCTAGAGCAATTTAACTCTGAGGCTATC 1098
DB 2 ATCTCGAACTGTCACGACGACGATGMAAGCTAAAGCAATTTAACTCTGAGGCTATC 61
QY 1099 AACAGCAGATAG-AGTGTGACGACTGATCATGACGACGAGGCTTTTGTCTAAAGAAAGTA 1157
DB 62 AACAGCAGATAGAAATGAAAGACTGATCATCAAGCAG- GCTTTTGTCTAAAGAAAGTA 120
QY 1158 ACAGNAGAGAGTACAAACACACTGTTCTAGACCCAAAATGAGCTGATTTGATAT-G 1216
DB 121 ACAGAGAGAGTACAAACACACTGTTCAAGACCCCAATTAATGAACTGAAATTTGATATGG 180
QY 1217 GCTAGCTGAGNAGAAATCATCATCAAGCAGCTTAAATGAGGCTGCTTATTTAGCC 1276
DB 181 GATAGCCTGAGAGAGAAATCATCAAGCAGCTTAAATGAGGCTGCTTATTTAGCC 240
QY 1277 NTGGGTCACTTCCAACTGCTGTATGAGCTCTG--CATGCTGACACACACGAGGGC 1334
DB 241 TTGGGTCACTTCCAACTGCTGTATGAGCTCTGAGCTCTGAGCTGACACACACGAGGGC 300
QY 1335 TTGCTAAGTGAGCAAAACCTGTTGAGAGAGACCTTAAAGCCAT 1378
DB 301 TTGCTAAGTGAGCAAAACCTGTTGAGAGAGACCTTAAAGCCAT 344

RESULT 14
US-10-043-487-126
;/ Sequence 126, Application US/10043487
;/ Publication No. US2003055220A1
;/ GENERAL INFORMATION:
;/ APPLICANT: HYBRIGENICS
;/ APPLICANT: PIERRE, UEGRAIN
;/ TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
;/ FILE REFERENCE: B4778A
;/ CURRENT APPLICATION NUMBER: US/10/043,487
;/ PRIOR FILING DATE: 2002-04-30
;/ PRIOR APPLICATION NUMBER: US 60/261,130
;/ NUMBER OF SEQ ID NOS: 561
;/ SOFTWARE: PatentIn version 3.1
;/ SEQ ID NO 126
;/ LENGTH: 1721
;/ TYPE: DNA
;/ ORGANISM: Shigella flexneri
US-10-043-487-126

Query Match 11.6%; Score 235.6; DB 14; Length 1721;
Best Local Similarity 66.6%; Pred. No. 1.1e-50;
Matches 337; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 1316 GCTGACACACACCGAGGCTTGTCTAAGTAGAGCAAACTGTTGAGAGAGACCTTAAAGC 1375
DB 1 GCTGACACACACCGAGAGAGTGTCTAAGTAGAGCAAACTGTTGAGAGAGACCTTAAAGT 60
QY 1376 CATTGAATTTGAATCTGCGAAGCATCATGTCTCCAAATGATGATTTAGCCCATCATGTC 1435

DB 61 CATTGAATTTGAATCTGCGAAGCATCATGTCTCCAAATGATGATTTAGCCCATCATGTC 120
QY 1436 CACAGTGAAGCGCTTAATTAAGCAGGAATATCTAATTTGAATCAAGTGACGAGAGTA 1495
DB 121 CACAGTGAAGCGCTTAATTAAGCAGGAATATCTAATTTGAATCAAGTGACGAGAGTA 180
QY 1496 AGCAAGCAACTTCAGAACAGCTAGAGGTTTAAATCAACGCTGCAAAATGTTTTGGA 1555
DB 181 TGCAGCAGCTTAAGAGAGCGTTTGAAGCATTGAACCATGAACCATGTGGAGATGATTTAC 240
QY 1556 AAAACAGAACAG 1615
DB 241 GAAAACAG 300
QY 1616 TGGCGAAATTTGAGATTTGACAGAGTGGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 1675
DB 301 CAGTGAATTTGAGATTTGACAGAGTGGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1676 TAAACCGCTGGAGGTTTACCGGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1735
DB 361 TAAAGCCAG 420
QY 1736 CTGTGCTGCTTGAAGCTTAAAGAGAAACATTAATTAAGTCTGATGACAGAAAGCCAGCA 1795
DB 421 CTATTCAGAGCTGAAG 480
QY 1796 GATGCTTGAAGATGAG 1821
DB 481 CAGCTTGAAG 506

RESULT 15
US-10-094-749-272
;/ Sequence 272, Application US/10094749
;/ Publication No. US20030219741A1
;/ GENERAL INFORMATION:
;/ APPLICANT: ISOGAL, TAKAO
;/ APPLICANT: SUGIYAMA, TOMOYASU
;/ APPLICANT: OTSUKI, TETSUJI
;/ APPLICANT: WAKAMATSU, AI
;/ APPLICANT: SATO, HIROYUKI
;/ APPLICANT: ISHII, SHIZUKO
;/ APPLICANT: YAMAMOTO, JUN-ICHI
;/ APPLICANT: ISONO, YUUKO
;/ APPLICANT: HIO, YURI
;/ APPLICANT: OTSUKA, KAORU
;/ APPLICANT: NAGAI, KEIICHI
;/ APPLICANT: IRIE, RYOTARO
;/ APPLICANT: TAMECHIKI, ICHIRO
;/ APPLICANT: SEKI, NAOHITO
;/ APPLICANT: YOSHIKAWA, TSUTOMU
;/ APPLICANT: OTSUKA, MOTOMYUKI
;/ APPLICANT: NAGAHARI, KENJI
;/ APPLICANT: MASUHO, YASUHIKO
;/ TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;/ FILE REFERENCE: 08435/0160
;/ CURRENT APPLICATION NUMBER: US/10/094,749
;/ PRIOR FILING DATE: 2002-03-12
;/ PRIOR APPLICATION NUMBER: 60/350,435
;/ PRIOR FILING DATE: 2002-01-24
;/ PRIOR APPLICATION NUMBER: JP 2001-328381
;/ PRIOR FILING DATE: 2001-09-14
;/ NUMBER OF SEQ ID NOS: 3381
;/ SOFTWARE: PatentIn Ver. 2.1
;/ SEQ ID NO 272
;/ LENGTH: 3684
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
US-10-094-749-272

Query Match 8.8%; Score 177.8; DB 15; Length 3684;
Best Local Similarity 92.5%; Pred. No. 2.1e-35;
Matches 210; Conservative 0; Mismatches 12; Indels 5; Gaps 2;

```

QY      8 GCTGACAGAAACACATCATCTCTCAGC-TCCGCGCCGAGCCCTGGAATGATAC 66
Db      3458 GCTGACAGAAACACATCATCTCTCAGCTTCCGCCCCAGCCCTTGAATATGAAAC 3517
QY      67 TCTAAGGCAGCAGCAGAGAAACATCGGCACTGCGTAGTTGATAGCTGAACACAAGCC 126
Db      3518 TCTAAGGCAGCAGCAGAGAAACATCGGCACTGCGTAGTTGATAGCTGAACACAAGCC 3577
QY      127 TCATATGATAGATGAAACAACTCG---CCAAGTACTGAATGAGCCCTGGGGAAAG 182
Db      3578 TCATATGATAGATGAAACAACTCGGCGCACAGTTACTGGAATTGAGCCCTGGGGAAAG 3637
QY      183 CTTTCTATCCAGAGAGATATGTGGCAGCGACACCCCTTACAGTC 229
Db      3638 CTTTCTATCCAGAGAGATATGTGGCAGCGACACCCCTTACAGTC 3684

```

Search completed: January 12, 2005, 15:13:14
 Job time : 1090 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2005, 07:18:34 : Search time 193 Seconds

(without alignments)
7479.856 Million cell updates/sec

Title: US-10-089-887-46

Perfect score: 2031

Sequence: 1 ttcgcatgcgcagcagaaca.....tngagagctcccaacgcgc 2031

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: Issued Patents NA:*
- 2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	266.8	13.1	4495	US-09-620-312D-352	Sequence 352, App
2	250.2	12.3	278	US-09-513-999C-36133	Sequence 36133, A
3	93.8	4.6	1024	US-09-328-475C-9	Sequence 9, Appli
4	69.4	3.4	1613	US-08-938-830-28	Sequence 28, Appli
5	67	3.3	7218	US-08-232-463-14	Sequence 14, Appli
6	54.8	2.7	2311	US-09-720-317A-19	Sequence 19, Appli
7	52.8	2.6	2519	US-09-380-287A-9	Sequence 9, Appli
8	50.4	2.5	3472	US-09-873-737A-5	Sequence 5, Appli
9	50.2	2.5	1024	US-09-328-475C-37	Sequence 37, Appli
10	50	2.5	2365	US-09-249-687A-5	Sequence 5, Appli
11	50	2.5	2365	US-09-249-687A-18	Sequence 18, Appli
12	50	2.5	2365	US-09-363-316B-5	Sequence 5, Appli
13	50	2.5	2365	US-09-363-316B-23	Sequence 23, Appli
14	48.6	2.4	1493	US-08-340-820-24	Sequence 24, Appli
15	48.6	2.4	1493	US-08-593-535-24	Sequence 24, Appli
16	48.6	2.4	1638	US-09-591-095-26	Sequence 26, Appli
17	48	2.4	8607	US-10-204-708-72	Sequence 72, Appli
18	46.8	2.3	1459	US-09-537-654-3	Sequence 3, Appli
19	46.8	2.3	1663	US-09-464-535-43	Sequence 43, Appli
20	46.8	2.3	2527	US-09-244-805-29	Sequence 29, Appli
21	46.4	2.3	830	US-08-688-609-1	Sequence 1, Appli
22	46.4	2.3	830	US-09-002-832-1	Sequence 1, Appli
23	46.2	2.3	1024	US-09-328-475C-36	Sequence 36, Appli
24	46.2	2.3	1069	US-09-205-258-74	Sequence 74, Appli
25	46.2	2.3	2799	US-09-181-339-2	Sequence 2, Appli
26	46.2	2.3	3124	US-09-734-030-1	Sequence 1, Appli
27	46.2	2.3	3124	US-10-153-921-1	Sequence 1, Appli

28	45.8	2.3	997	US-09-907-794A-376	Sequence 376, App
29	45.8	2.3	997	US-09-905-125A-376	Sequence 376, App
30	45.8	2.3	997	US-09-902-775A-376	Sequence 376, App
31	45.8	2.3	997	US-09-906-700-376	Sequence 376, App
32	45.8	2.3	997	US-09-903-603A-376	Sequence 376, App
33	45.6	2.2	1361	US-09-489-847-64	Sequence 64, Appli
34	45.6	2.2	1674	US-10-140-002-453	Sequence 453, App
35	45.6	2.2	2394	US-09-800-729-13	Sequence 33, Appli
36	45.4	2.2	3396	US-08-989-299-1	Sequence 1, Appli
37	45.4	2.2	3396	US-10-158-847-141	Sequence 141, App
38	45.4	2.2	3396	US-09-407-427-1	Sequence 1, Appli
39	45.2	2.2	708	US-09-270-767-13081	Sequence 13081, A
40	45.2	2.2	1024	US-09-328-475C-35	Sequence 35, Appli
41	45.2	2.2	2447	US-09-014-969-14	Sequence 14, Appli
42	45	2.2	552	US-08-807-784B-8	Sequence 8, Appli
43	45	2.2	931	US-09-482-273-31	Sequence 31, Appli
44	45	2.2	1801	US-09-709-103-3	Sequence 3, Appli
45	45	2.2	1801	US-09-439-410A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-352

Sequence 352, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyen

APPLICANT: Zhao, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: PC_FL_genes Version 1.0

SEQ ID NO 352

LENGTH: 4495

TYPE: DNA

ORGANISM: Homo sapiens

FEATURES:

NAME/KEY: CDS

LOCATION: (307)..(3006)

US-09-620-312D-352

Query Match

Best Local Similarity 93.9%: Pred. No. 1.1e-64

Matches 277: Conservative 0: Mismatches 18: Indels 0: Gaps 0:

QY 1593 TTGGCCAGGCGCAAGGTTTCATGCGCAATTGAGATTTCAGCAGTGGCTGACTGAC 1652
Db 172 TTACTTACGCGCAAGGTTTCATGCGCAATTGAGATTTCAGCAGTGGCTGACTGAC 231

```
QY 1653 ACCGAGGTCATCTGTTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAG 1712
|||
Db 232 ACCGAGGTCATCTGTTGGCATCTAAACCGCTGGAGGTTTACCGGAAACAGCCAAAG 291
|||
QY 1713 CAGCTTAATCTCCATATGGAAGTCTGTGCTGCTCTTGAAGCTTAAAGAAACATTAAG 1772
|||
Db 292 CAGCTTAATCTCCATATGGAAGTCTGTGCTGCTCTTGAAGCTTAAAGAAACATTAAG 351
|||
QY 1773 AGCTGATGAGAAAGGCCAGAGATGCTGCAAGATGCCCAAAATCTGCAGAGCAAAAT 1832
|||
Db 352 AGCTGATGAGAAAGGCCAGAGATGCTGCAAGATGCCCAAAATCTGCAGAGCAAAAT 411
|||
QY 1833 ATTGACCAAGACATTAATTAATCTGAAAAAATGAAAAAATGAAAAAATGAAAAAAT 1887
|||
Db 412 ATTGACCAAGACATTAATTAATCTGAAAAAATGAAAAAATGAAAAAATGAAAAAAT 466
|||

RESULT 2
US-09-513-999C-36133
; Sequence 36133, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 36133
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 153
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 154
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 263
; OTHER INFORMATION: w=a or t
US-09-513-999C-36133

Query Match 12.3%; Score 250.2; DB 4; Length 278;
Best Local Similarity 98.2%; Pred. No. 1e-60;
Matches 271; Conservative 2; Mismatches 1; Indels 2; Gaps 2;
```

```
Db 241 AGACATAAGGGAAGAAATAGAGACTACAGAGG 276
|||

RESULT 3
US-09-328-475C-9
; Sequence 9, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Caroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-9

Query Match 4.6%; Score 93.8; DB 4; Length 1024;
Best Local Similarity 82.5%; Pred. No. 2.7e-16;
Matches 118; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
```

```
QY 1846 TAAATACCTGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAAT 1905
|||
Db 322 TAATTAAACCAAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAAT 381
|||
QY 1906 ATTACAGCGCGCTGTAATCTAGACCTGCCGCGCGCGCTGAGCCCTATAGTAGT 1965
|||
Db 382 ATTAAGCGCGCGCTGTAATCTAGACCTGCCGCGCGCGCTGAGCCCTATAGTAGT 440
|||
QY 1966 CGTATTAGATGGAATCACTAGT 1988
|||
Db 441 CGTATTAGATGGAATCACTAGT 463
|||

RESULT 4
US-08-938-830-28/C
; Sequence 28, Application US/08938830
; Patent No. 6040437
; GENERAL INFORMATION:
; APPLICANT: Laeky, Laurence A.
; APPLICANT: Dowbenko, Donald J.
; TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,830
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/798419
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P106BP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1613 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-938-830-28

Query Match 3.4%; Score 69.4; DB 3; Length 1613;
Best Local Similarity 87.9%; Pred. No. 2,7e-09;
Matches 87; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 1996 GAATTCGAAATTCAGCGCGCGCTGAATTCGAACTCCCGCGCGCGCGCTCGAGCC 1955
Db 164 GAAAGAAGAAATTCAGCGCGCGCTGAATTCGAACTCCCGCGCGCGCGCTCGAGCC 106
QY 1956 TATAGTAGTGTCTATTAGATGGAATCTACTGTGCGCGC 1994
Db 105 TATAGTAGTGTCTATTAGATGGAATCTACTGTGCGCGC 67

RESULT 5
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHERFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109

```

```

: TELEX 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZgt-F18
: US-08-232-463-14

Query Match 3.3%; Score 67; DB 1; Length 7218;
Best Local Similarity 7.1%; Pred. No. 3e-09;
Matches 31; Conservative 231; Mismatches 173; Indels 0; Gaps 0;

Oy 1492 AGAAGACAGCAACCTTCAGAACAGGTAGAGTTTAAATCAACGCTGGCAAAATGTTT 1551
Db 1457 AAAGAGATAGAAAGATTGTTGACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
Oy 1552 TGGAAAAAACGAAACAAAGAGACAGACGCTGATGTCCTTCCGACAGCCAAAGGT 1611
Db 1397 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1338
Oy 1612 TCATGCGCAATTAGAGATTTCAGAGAGCGCTGATGACACAGACGTCATCTGTTGG 1671
Db 1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278
Oy 1672 CATCTAAACCGCTGGAGTTTACCGGAAACAGCAAGACGACCTTAATGTCATATGG 1731
Db 1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218
Oy 1732 AAGTCGTGCTGCTTGAAGCTAAGAGAAACATATAAGAGTCGTAGTCAGAAAGCC 1791
Db 1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158
Oy 1792 AGCAGATGCTGCAAGATGCCCAAAATCTGCAGAGACAAATATGACCAAGACTTAATA 1851
Db 1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1098
Oy 1852 ACTTGAATAAAAAAAAAAAAAAAAAAAAAAAAAANACNGCGCTGAATTCAGAAATTCAG 1911
Db 1097 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1038
Oy 1912 CGGCGCGCTGAATTC 1926
Db 1037 CTCGAATTAATTC 1023

RESULT 6
US-09-720-317A-19
: Sequence 19, Application US/09720317A
: Patent No. 6696292
: GENERAL INFORMATION:
: APPLICANT: Stephen M. Allen
: APPLICANT: Saverio C. Falco
: APPLICANT: Catherine J. Thorpe
: TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
: FILE REFERENCE: BB-1167
: CURRENT APPLICATION NUMBER: US/09/720,317A
: CURRENT FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: 60/092,833
: PRIOR FILING DATE: 14-07-1998
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 19
: LENGTH: 2311
: TYPE: DNA
: ORGANISM: Trifolium aestivum
US-09-720-317A-19

Query Match 2.7%; Score 54.8; DB 4; Length 2311;
Best Local Similarity 58.3%; Pred. No. 4.3e-05;
Matches 95; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

[illegible]

RESULT 7
US-09-380-287A-9
; Sequence 9, Application US/09380287A
Defendant's 0710171

1 APPLICANT: NAKAMURA, YUSUKE
 2 APPLICANT: TANAKA, TOSHIHIRO
 3 APPLICANT: TSUKADA, SHUICHI
 4 TITLE OF INVENTION: PHYSIOLOGICALLY ACTIVE PROTEIN ORIGINATING IN MAMMAL
 5 FILE REFERENCE:
 6 CURRENT APPLICATION NUMBER: US/09/380,287A
 7 CURRENT FILING DATE: 1999-11-16
 8 PRIOR APPLICATION NUMBER: JP 1997-62259
 9 PRIOR FILING DATE: 1997-02-28
 10 PRIOR APPLICATION NUMBER: JP 1998-62263
 11 PRIOR FILING DATE: 1998-02-25
 12 NUMBER OF SEQ ID NOS: 42
 13 SOFTWARE: PatentIn Ver. 2.1

```

Seq ID NO 9
LENGTH: 2519
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1785)
NAME/KEY: misc feature
LOCATION: (289)..(291)
OTHER INFORMATION: Xaa is Phe or Leu
NAME/KEY: misc_feature
LOCATION: (1015)..(1017)
OTHER INFORMATION: Xaa is Lys or Arg
NAME/KEY: misc feature
LOCATION: (1492)..(1494)
OTHER INFORMATION: n is a o r g or c or t
OTHER INFORMATION: Xaa is Gln or Lys or Glu
NAME/KEY: misc feature
LOCATION: (1540)..(1542), (1562)..(1584), (1653)..(1695)
OTHER INFORMATION: Xaa is Phe or Ser
NAME/KEY: misc_feature
LOCATION: (1609)..(1611)
OTHER INFORMATION: Xaa is Gln or Arg
NAME/KEY: misc feature
LOCATION: 2493..2516
OTHER INFORMATION: n is a o r g or c or t
OS-09-380-287A-9

```

	Query Match	Best Local Similarity	Score	DB 4	Length	2519;
	Matches	71; Conservative	83.5%; Pred. No. 0.00016;	0; Mismatches	10; Indels	4; Gaps
QY	1847	AAATPACTGTGAAAAAAAAAAAAAAAAA	ACAGCCGCTGAATTC	----	T	19022
Db	2435	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AGCGCGCTGAATCTAGT			24944
QY	1903	AGAAATCAGCGCCGCTGAATTC				1927
Db	2495	AGAAATCAGCGCCGCTGAATTC				2519

RESULT 8
ITS-09-A7

US-09-873-737A-5/C
Sequence 5, Application US/09873737A
Patent No. 6723554
GENERAL INFORMATION:
APPLICANT: Duke University
APPLICANT: Lin, Hailan
TITLE OF INVENTION: PURIFIED AND ISOLATED p14¹ FAMILY GENES AND GENE
TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
FILE REFERENCE: Attorney Docket No. 6723554 180-104/2
CURRENT APPLICATION NUMBER: US/09/873-737A

```

1 PRIOR APPLICATION NUMBER: PCT/US99/28766
2
3 PRIOR FILING DATE: 1999-12-03
4
5 PRIOR APPLICATION NUMBER: 60/110,901
6
7 PRIOR FILING DATE: 1998-12-04
8
9 NUMBER OF SEQ ID NOS: 21
10
11 SOFTWARE: PatentIn Ver. 2.1
12
13 SEQ ID NO:

```

```

? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (167)..(2749)
? NAME/KEY: misc feature
? LOCATION: (1392)
? OTHER INFORMATION: n=t or a, Xaa=Leu or Ile
? NAME/KEY: misc feature
? LOCATION: (1073)
? OTHER INFORMATION: n=c or a, Xaa=Leu or Ile
? NAME/KEY: misc feature
? LOCATION: (2365)
? OTHER INFORMATION: n=c or a, Xaa=Leu or Ile
?
? OS-09-873-737A-5

```

Query Match	2.5% ;	Score 50.4 ;	DB 4 ;	Length 3472 ;
Best Local Similarity	77.1% ;	Pred. No. 0.0094 ;		
Matches	74 ;	Conservative 0 ;	Mismatches 21 ;	Indels 1 ;
QY	1913	GGCGCGCAATTCTAGACCTGCCCGGGCGCGGCTCGAGCCCATATGATGTCGTATTA	1972	
DB	126	GGTCCTTGCACTTCAGACCTGCCCGGGCGGC-CGCTCGAGCCCATATGTCGTATTA	68	
QY	1973	GGATGGAATCACTAGTCGCGCGCCTCGAGGTCGACC	2008	
DB	67	GGATGGAATGAAATTCCTCGCGCGCGCCATGCGGCC	32	

```

RESULT 9
US-09-328-475C-37/c
Sequence 37, Application US/09328475C
Patent No. 6476207
GENERAL INFORMATION:
APPLICANT: Zhang, Jimmy
APPLICANT: Astel, Jon H.
APPLICANT: Carroll III, Eddie
APPLICANT: Endege, Wilson O.
APPLICANT: Ford, Donna M.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
APPLICANT: Steinmann, Kathleen E.
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
FILE REFERENCE: 1532.002/200130.463
CURRENT APPLICATION NUMBER: US/09/328,475C
NUMBER OF SEQ ID NOS: 341
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 37
LENGTH: 1024
TYPE: DNA
ORGANISM: Homo Sapien

```


FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1024)
OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-37

Query Match 2.5%; Score 50.2; DB 4; Length 1024;
Best Local Similarity 65.6%; Pred. No. 0.00052;
Matches 103; Conservative 0; Mismatches 50; Indels 4; Gaps 2;

QY 1832 TATTGACCAACATTAATTAATTGAAAAA
DB 200 TTTTAAAGAAAAAGCAAAATTAACAAAAAGCCCGCTGTA
QY 1892 CGCTGAATTGAGATTCAGCGCGCGCTGAATTCTAGACCTGCGCGCGCGCGCTGCA 1951
DB 140 AAAAAAAAAAAAAAAAAAGCGCGCGCTGAATTC---ACCTGCGCGCGCGCGC-CGCTGCA 85
QY 1952 GCCCTATAGTAGTCGATTAGATGGAATCACTAGT 1988
DB 84 GCCCTATAGTAGTCGATTAGATGGAAGCGCGAAT 48

RESULT 10

US-09-249-697A-5/C
Sequence 5, Application US/09249697A

PATENT NO. 6392018
GENERAL INFORMATION:
APPLICANT: Ford, John
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
FILE REFERENCE: 24011-727
CURRENT APPLICATION NUMBER: US/09/249,697A
CURRENT FILING DATE: 1999-02-12
PRIORITY APPLICATION NUMBER: US 08/968,800
PRIORITY FILING DATE: 1997-11-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 2365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)...(1866)
US-09-249-697A-5

Query Match 2.5%; Score 50; DB 3; Length 2365;
Best Local Similarity 98.4%; Pred. No. 0.00098;
Matches 61; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1927 AGACCTGCGCGCGCGCGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 1986
DB 61 AGACCTGCGCGCGCGCGC-CGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 3

QY 1987 GT 1988
DB 2 GT 1

RESULT 11
US-09-249-697A-18/C
Sequence 18, Application US/09249697A
PATENT NO. 6392018
GENERAL INFORMATION:
APPLICANT: Ford, John
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
FILE REFERENCE: 24011-727
CURRENT APPLICATION NUMBER: US/09/249,697A
CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 08/968,800
PRIOR FILING DATE: 1997-11-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18
LENGTH: 2365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)...(1863)
US-09-249-697A-18

Query Match 2.5%; Score 50; DB 3; Length 2365;
Best Local Similarity 98.4%; Pred. No. 0.00098;
Matches 61; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1927 AGACCTGCGCGCGCGCGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 1986
DB 61 AGACCTGCGCGCGCGCGC-CGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 3
QY 1987 GT 1988
DB 2 GT 1

RESULT 12

US-09-363-316B-5/C
Sequence 5, Application US/09363316B

PATENT NO. 6392019
GENERAL INFORMATION:
APPLICANT: Ford, John
TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
FILE REFERENCE: 28110/35852
CURRENT APPLICATION NUMBER: US/09/363,316B
CURRENT FILING DATE: 1999-07-28
PRIORITY APPLICATION NUMBER: US 09/249,697
PRIORITY FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 08/968,800
PRIOR FILING DATE: 1997-11-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 2365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)...(1866)
NAME/KEY: misc feature
LOCATION: (357)
OTHER INFORMATION: Xaa = any amino acid
US-09-363-316B-5

Query Match 2.5%; Score 50; DB 3; Length 2365;
Best Local Similarity 98.4%; Pred. No. 0.00098;
Matches 61; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1927 AGACCTGCGCGCGCGCGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 1986
DB 61 AGACCTGCGCGCGCGCGC-CGCTCGAGCCCTATAGTACGATTAGGATGAATCACTA 3
QY 1987 GT 1988
DB 2 GT 1

RESULT 13
US-09-363-316B-23/C
Sequence 23, Application US/09363316B
PATENT NO. 6392019
GENERAL INFORMATION:

RESULT 14
US-08-340-820-24
; Sequence 24: Application US/08340820
; Patent No. 5512460
; GENERAL INFORMATION:
; APPLICANT: NARUO, Ken-ichi
; APPLICANT: SEKO, Chisako
; APPLICANT: KUROKAWA, Tautomu
; APPLICANT: KONDO, Tatsuya
; TITLE OF INVENTION: GLIA ACTIVATING FACTOR AND ITS
; TITLE OF INVENTION: PRODUCTION
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,820
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,713
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CONLIN, David G.
; REGISTRATION NUMBER: 27026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE

RESULT 15
 US-08-593-535--24
 Sequence 24, Application US/08593535
 Patent No. 5622928
 GENERAL INFORMATION:
 APPLICANT: NARUO, Ken-ichi
 APPLICANT: SEKO, Chiako
 APPLICANT: KUROKAWA, Tsutomu
 APPLICANT: KONDO, Tatsuya
 TITLE OF INVENTION: GLIA ACTIVATING FACTOR AND ITS
 TITLE OF INVENTION: PRODUCTION
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
 ADDRESSEE: CUSHMAN
 STREET: 130 Water Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/593,535
 FILING DATE: 24-JAN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/835,713
 FILING DATE: 12-FEB-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: CONLIN, David G.
 REGISTRATION NUMBER: 27026
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)523-3400

TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
HAPLOTYPE: 2n
TISSUE TYPE: skin
CELL TYPE: fibroblast
IMMEDIATE SOURCE:
LIBRARY: Human foreskin CDNA library
CLONE: pGAF1
US-08-593-535-24

Query Match 2.4%; Score 48.6; DB 1; Length 1493;
Best Local Similarity 58.3%; Pred. No. 0.0019;
Matches 84; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1745 CTTGAAGCTAAGAGAAATATTAAGTCTGATGCAGAAAGCCACGATGCTTGC 1804
DB 1349 CTTAGTAAAAATATAAAAAATATAAAAAATATAAAAAAGTTAAATTTATTATAG 1408
QY 1805 AAGATGCCCAAAATCTGCAGAGACAATAATTGACCAAGACATAAATACTGAAGAAAAA 1864
DB 1409 AAATTCAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1468
QY 1865 AAAAAAAAAAAAAAAAAANACA 1888
DB 1469 AAAAAAAAAAAAAAAAAAAAAA 1492

Search completed: January 12, 2005, 12:52:48
Job time : 195 secs

This Page Blank (uspto)